Proceedings of the Indian Academy of Sciences

(Mathematical Sciences)

Volume 95, 1986

CONTENTS

Proc. Indian Acad. Sci. (Math. Sci.), Vol. 95, No. 1, September 1986, pp. 1–12.

Axisymmetric melting of a long cylinder due to an infinite flux

S C GUPTA
Department of Applied Mathematics, Indian Institute of Science, Bangalore 560 012, India

MS received 4 July 1985; revised 3 March 1986

Abstract. By employing a new embedding technique, a short-time analytical solution for the axisymmetric melting of a long cylinder due to an infinite flux is presented in this paper. The sufficient condition for starting the instantaneous melting of the cylinder has been derived. The melt is removed as soon as it is formed. The method of solution is simple and straightforward and consists of assuming fictitious initial temperature for some fictitious extension of the actual region.

Keywords. Embedding technique; moving boundary problems; ablation; isotherm condition; heat-balance condition.

1. Introduction

The melting problem is essentially a two-phase heat transfer problem in which both liquid and solid phases are present. The problem considered in this paper is a one-phase problem in which the melt is removed as soon as it is formed. This problem is known as the ablation problem and may arise in many physical situations such as ablation of heat shields during re-entry of spacecraft into the earth's atmosphere, laser heating, oxygen diffusion with absorption etc.

Melting problems can be put under a more general class of problems known as moving boundary problems (MBPs). The solidification problem is mathematically analogous to the melting problem and there is no need to discuss the solidification problem separately and any reference to melting problems includes solidification problems as well. A typical feature of the MBPs is that, apart from the fixed boundary of the region under consideration, there is also a boundary inside the region which is unknown and is moving with time. Some boundary conditions are to be satisfied on this moving boundary. Depending on the portion of the fixed boundary on which the melting initiates, Sikarskie and Boley [20] have divided these melting problems into three classes. In class I problems, the melting starts simultaneously at all points on the fixed boundary. In class II and III problems, the melting starts over a portion of the fixed boundary and at a single point on the boundary respectively. This paper is concerned with a class I problem.

The unknown nature of the moving boundary and the boundary conditions to be satisfied on this moving boundary make the problems extremely complicated and their exact analytical solutions valid for total melting time and varied shapes are

rather unconceivable. Class I problems are easier to tackle as compared to class II and III problems as in the latter case the problem has to be at least two-dimensional and the boundary conditions on the fixed boundary are very difficult to manage.

For one dimensional class I melting problems, some exact similarity solutions exist in the literature such as Neumann solution [5], the series solution by Tao [21], the solution by Özisik and Uzzell [16] etc. Some more references can be found in the above mentioned references. These solutions pertain to very simple situations and geometries. Perturbation methods [17], approximate methods [10] and numerical methods [6, 7] have also been employed to study the melting problems. Fairly good account of the studies on MBPs can be found in [2, 13, 15]. Except numerical methods, sufficient applications of other methods to multi-dimensional problems are not available. Some of the numerical methods can be found in [9]. The utility of short-time analytical solutions in numerical schemes has been amply demonstrated in a recent work by Schulze *et al* [19] in which a practical problem of ingot solidification was considered. Short-time solutions can be used for starting some of the numerical schemes such as Murray and Landis scheme [14] and enthalpy scheme [8] and can also be used for checking the validity of such numerical schemes in which modifications are attempted.

Short time solutions require some special techniques. Boley *et al* [4] (which contains many references of the works of Boley and co-workers) have contributed substantially in developing short-time solutions of ablation and solidification problems pertaining to one-, two-, and three-dimensional regions. Recently Gupta and Lahiri [12] and Gupta [11] have studied short-time solutions of solidification in cylindrical mold (one-dimensional) and semi-infinite mold (two-dimensional), respectively. In [3], Boley developed an embedding technique which consists of prescribing fictitious flux in the case of ablation problems and fictitious boundary flux (or fictitious boundary temperature) together with suitable fictitious initial temperature in the case of solidification problems. The solid and liquid regions are embedded in the region originally occupied by the melt. The temperature solution in Boley's technique is written with the help of Duhamel's theorem [5] and the moving boundary and other unknowns are determined by comparing different powers of time variable on both sides of some integro-differential equations. In the present method, the temperature solution is written with the help of the source solution and the solid region is first embedded in the original region occupied by the solid and then the original solid region is extended fictitiously and some fictitious initial temperature is prescribed in the fictitious extension. The unknowns can be evaluated by differentiating the equations with respect to the time variable and taking appropriate limits. This yields considerable simplification and the method of solution becomes straightforward.

In the present melting problem, the prescribed flux is infinite at time $t = 0$ and some remarks about taking this type of singular flux will be in order. In many physical situations the melting takes place instantaneously i.e. there is no preheating of the solid before melting and the melting starts as soon as the flux is applied. If the boundary of the solid at which flux is applied is at a temperature

lower than the melting temperature then for instantaneous melting it is necessary (but not sufficient) that the applied flux is infinite and has singularity of the type $t^{-1/2}$. This type of singularity in the context of heat transfer coefficient has been experimentally observed earlier by Ruddle [18] and Tiller [22].

2. Problem formulation

A long cylindrical solid occupies the region $0 \le r \le R_0$, $|z| < \infty$ at time $t = 0$ (axisymmetric problem in which r and z are cylindrical polar coordinates). This solid is heated and the prescribed flux $Q(z, t)$ on the surface $r = R_0$ of the cylinder is a known quantity which is infinite at $t = 0$. $f(r, z)$ is the known initial temperature of the solid. It will be assumed that the melting starts instantaneously all over the boundary, $R = R_0$ (which happens, provided a sufficient condition is satisfied which will be derived later) and melt is removed as soon as it is formed. With time, the melting will progress along the interior till the whole of the solid is melted. The equation of the melting front may be written in the form

$$r = s(z, t). \tag{1}$$

The quantities of interest in the present study are the melting front and the temperature in the solid. If the melting does not start at $t = 0$ but starts at the time $t = t_m$, $t_m > 0$, then the problem in general will be a mixed problem of class II and III type which cannot be solved by the present technique and for which absolutely no reference is available in the literature.

The following dimensionless differential equation, initial and boundary conditions are to be satisfied:

$$2\frac{\partial T}{\partial V} = V\left(\frac{\partial^2 T}{\partial R^2} + \frac{1}{R}\frac{\partial T}{\partial R} + \frac{\partial^2 T}{\partial Z^2}\right) \tag{2}$$

$$0 \le R < S(Z, V),\ |Z| < \infty,\ V > 0,$$

$$T(R, Z, 0) = f(R, Z); \quad \left.\frac{\partial T}{\partial R}\right|_{R=0} = 0, \tag{3a, 3b}$$

$$T(R, Z, V)\big|_{R=S(Z,V)} = 1, \tag{4}$$

$$\left\{1 + \left(\frac{\partial S}{\partial Z}\right)^2\right\} \left.\frac{\partial T}{\partial R}\right|_{R=S} = Q(Z, V) + \frac{2\lambda}{V}\frac{\partial S}{\partial V}, \tag{5}$$

$$S(Z, V)\big|_{V=0} = 1. \tag{6}$$

$T(R, Z, V)$ is the dimensionless temperature. All the temperatures have been made dimensionless with the help of the melting temperature T_m which is taken as unique. Equation (4) is the isotherm condition and (5) is the heat balance condition in which the resultant heat flux vector is along the direction of R. The following dimensionless variables have been used in writing equations (2)–(6).

$$Z = z/R_0,\ R = r/R_0,\ V = 2(kt/R_0^2)^{1/2}, \tag{7}$$

$$\lambda = l/CT_m,\ f(R,\ Z) = f(r,\ z)/T_m, \tag{8}$$

$$Q(Z,\ V) = Q(z,\ t).\ R_0/KT_m,\ S(Z,\ V) = s(z,\ t)/R_0, \tag{9}$$

R_0 is the radius of the cylinder, K is the thermal conductivity, k is the diffusivity, l is the latent heat and C is the specific heat of the solid. Thermal properties are taken to be constants.

3. Solution

The solution of (2) with (3) can be written as

$$T(R, Z, V) = \frac{2}{\pi^{1/2}V^3}\left[\int_{-\infty}^{\infty}\int_0^1 H(p, q)f_2(p, q)\,\mathrm{d}p\mathrm{d}q + \int_{-\infty}^{\infty}\int_1^{\infty} H(p,\ q)f_2(p,\ q)\,\mathrm{d}p\mathrm{d}q\right], \tag{10}$$

$$0 \le R < \infty,\ |Z| < \infty,\ V > 0,$$

$$H(p,\ q) = p\exp[-\{R^2+p^2+(Z-q)^2\}/V^2]I_0(2Rp/V^2). \tag{11}$$

$I_0(x)$ is the modified Bessel function of the first kind of order zero.

It can be easily checked that $T(R,\ Z,\ V)$ in (10) satisfies (2) and (3). $f_2(R,\ Z)$ is the unknown initial temperature in the fictitious extension $1 \le R < \infty$, $|Z| < \infty$ of the original region. Mathematically, for the determination of two unknowns, namely, $S(Z,\ V)$ and $f_2(R,\ Z)$, there are two conditions (4) and (5) to be satisfied. We shall first formally obtain the short time analytical solution.

For large values of the argument, the following asymptotic series expansion can be used for $I_0(x)$ [5]

$$I_0(x) = \frac{\exp(x)}{(2\pi x)^{1/2}}\left\{1 + \frac{1}{8x} + \frac{9}{128x^2} + \ldots.\right\}. \tag{12}$$

Firstly in (10), $I_0(x)$ is replaced by the first two terms of the asymptotic series in (12) and then (10) is substituted in (4) and (5). After making suitable substitutions the following equations are obtained

$$\pi S^{3/2} = \int_{\infty}^{-\infty}\int_{(S/V)}^{(S-1)/V} D_1(S,\ V,\ p)f_1(S-Vp,\ Z-Vq) \times\exp\{-(p^2+q^2)\}\,\mathrm{d}p\mathrm{d}q + \int_{\infty}^{-\infty}\int_{(S-1)/V}^{-\infty}\{D_1(S,\ V,\ p)/(S-Vp)^{1/2}\}\,f_2(S-Vp,\ Z-Vq)\exp\{-(p^2+q^2)\}\,\mathrm{d}p\mathrm{d}q, \tag{13}$$

$$\left\{1+\left(\frac{\partial S}{\partial Z}\right)^2\right\}\left[\int_{\infty}^{-\infty}\int_{(S/V)}^{(S-1)V} D_2(S,\ V,\ p)f_1(S-Vp,\ Z-Vq)\right.$$

$$\times \exp\{-(p^2+q^2)\}\, dp dq + \int_{\infty}^{-\infty} \int_{(S-1)/V}^{-\infty} \{D_2(S, V, p)/ (S-Vp)^{1/2}\} f_2(S-Vp, Z-Vq) \exp\{-(p^2+q^2)\}\, dp dq]$$

$$= \pi V Q(Z, V) + 2\pi\lambda \frac{\partial S}{\partial V}, \tag{14}$$

where

$$D_1 = S(S-Vp) + V^2/16, \tag{15}$$

$$D_2 = 2p\{S^2(S-Vp)\} + V^2 S/16\} + VS(S-Vp)/2 + 3V^3/32, \tag{16}$$

$$f_1(R, Z) = f(R, Z)/R^{1/2}. \tag{17}$$

The following series expansions for the known and unknown functions will be assumed.

$$f_1(R, Z) = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \frac{(R-1)^m Z^n}{\lfloor m \, \lfloor n} \left. \frac{\partial^{m+n}(f_2)}{\partial R^m \partial Z^n} \right|_{\substack{R=1 \\ Z=0}},$$
$$0 < R \leq 1, |Z| < \infty, \tag{18}$$

$$f_2(R, Z) = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \frac{(R-1)^m Z^n}{\lfloor m \, \lfloor n} \left. \frac{\partial^{m+n}(f_2)}{\partial R^m \partial Z^n} \right|_{\substack{R=1 \\ Z=0}},$$
$$1 < R < \infty, \; |Z| < \infty, \tag{19}$$

$$Q(Z, V) = \sum_{n=0}^{\infty} Q_n(Z) V^{n-1}, \; V > 0, \; |Z| < \infty, \tag{20}$$

$$S(Z, V) = 1 + \sum_{n=1}^{\infty} A_n(Z) V^n, \; V > 0, \; |Z| < \infty. \tag{21}$$

In (18) and (19), the functions have been expanded in powers of $(R-1)$ as the melting initiates at $R=1$ and in short time solutions we are interested in the behaviour of these functions around $R=1$. In (20), the flux is infinite at $V=0$ and mathematically the solution can be determined only if singularity at $V=0$ is of the type $1/V$. Equation (20) is motivated by the Neumann solution. In order to obtain the unknowns, the series expansions given above are substituted in (13) and (14) and then the limits $V \to 0+$ of these equations (13) and (14) are taken. Two equations in the two unknowns $A_1(Z)$ and $f_2(1, Z)$ are obtained which when solved give a unique solution. Equations (13) and (14) are then differentiated once with respect to V and limits $V \to 0+$ of these equations are taken. Once again two equations in the two unknowns $\partial f_2/\partial R\,|_{R=1}$ and $A_2(Z)$ are obtained which when solved give a unique solution. This process of higher order differentiations and limits $V \to 0+$ can be continued till the desired coefficient in the moving boundary is obtained. However, after few differentiations the algebra becomes extremely lengthy as with each higher order differentiation the number of terms goes on increasing. It may be remarked here that the limits of integrations as well as

integrands are functions of V and this should be taken care while differentiating the integrals and taking limits $V \to 0+$. Equation (13) determines the derivatives of the unknown function $f_2(R, Z)$ and (14) determines the unknown coefficients of the moving boundary.

Two relevant points may be pointed out here. Firstly, it has been assumed a priori that the unknown functions are sufficiently smooth so that operations of differentiations and limits $V \to 0+$ are valid. This cannot be proved as none of the unknown series can be determined completely. This sort of situation arises in many physical problems and what is done in such situations is to check whether the final outcome is correct or not and this has been done in the present work with the help of some analytical and numerical checks. Secondly, if the unknowns were not determined uniquely, then the method of solution would have failed.

The coefficients $A_1(Z)$, $A_2(Z)$ and $A_3(Z)$ in the moving boundary are given below. $A_1(Z)$ is the root of the following transcendental equation

$$2(\pi)^{1/2}\lambda A_1(Z)+\pi^{1/2}Q_0(Z)=\exp(-A_1^2)\,\{f_2(1, Z)-f_1(1, Z)\}, \tag{22}$$

$$f_2(1, Z)=\{2-\mathrm{erfc}(A_1)f_1(1, Z)\}/(1+\mathrm{erf}\,A_1). \tag{23}$$

If

$$(\pi)^{1/2}Q_0(Z)>2\{1-f_1(1, Z)\} \tag{24}$$

then $A_1(Z)<0$ for all Z, which implies that melting starts instantaneously and simultaneously all over the surface $R=1$ at $V=0$.

$$\begin{aligned}16\pi^{1/2}\lambda A_2(Z) &= -4\pi^{1/2}Q_1(Z)-20\pi^{1/2}A_1^2\\ &\quad +2\pi^{1/2}\left.\frac{\partial f_1}{\partial R}\right|_{R=1}\mathrm{erfc}\,A_1-\{8A_1\exp(-A_1^2)\\ &\quad -\pi^{1/2}\,\mathrm{erfc}\,A_1\}f_1(1, Z)+8A_1 f_2(1, Z)\exp(-A_1^2)\\ &\quad +2\pi^{1/2}(1+\mathrm{erf}\,A_1)\left.\frac{\partial f_2}{\partial R}\right|_{R=1},\end{aligned} \tag{25}$$

$$\begin{aligned}&\{2\exp(-A_1^2)+\pi^{1/2}A_1(1+\mathrm{erf}\,A_1)\}\left.\frac{\partial f_2}{\partial R}\right|_{R=1}\\ &\qquad =6\pi^{1/2}A_1+2\pi^{1/2}\mathrm{ierfc}\,(A_1)\left.\frac{\partial f_1}{\partial R}\right|_{R=1}\\ &+2(1+2A_2)\exp(-A_1^2)\,\{f_1(1, Z)-f_2(1, Z)\}+\{\exp(-A_1^2)\\ &-3\pi^{1/2}A_1(1+\mathrm{erf}\,A_1)\}f_2(1, Z)-4\pi^{1/2}A_1\,\mathrm{erfc}\,(A_1)f_1(1, Z),\end{aligned} \tag{26}$$

$$\begin{aligned}&96\pi^{1/2}\lambda A_3(Z)=-16\pi^{1/2}Q_2(Z)-240\pi^{1/2}\lambda A_1A_2\\ &-30\pi^{1/2}\lambda A_1^3+\{-16A_1'^2+32A_1A_3+8A_2+16A_2^2\\ &+64A_1^2A_2-32A_1^2A_2^2\}\exp(-A_1^2)\end{aligned}$$

$$\times\{f_1(1,Z)-f_2(1,Z)\}+8\{\pi^{1/2}A_1\operatorname{erfc}A_1$$
$$+(3-6A_2-2A_1^2)\exp(-A_1^2)\}f_1(1,Z)+4\{7\pi^{1/2}A_1\operatorname{erfc}A_1$$
$$-(3+4A_2)\exp(-A_1^2)\}\left.\frac{\partial f_1}{\partial R}\right|_{R=1}-8\pi^{1/2}\left.\frac{\partial^2 f_1}{\partial R^2}\right|_{R=1}\operatorname{ierfc}A_1$$
$$+(20A_2+16A_1^2-3)f_2(1,Z)\exp(-A_1^2)+4(1+4A_2)\exp(-A_1^2)$$
$$+5\pi^{1/2}A_1(1+\operatorname{erf}A_1)\left.\frac{\partial f_2}{\partial R}\right|_{R=1}+8\{\exp(-A_1^2)+\pi^{1/2}A_1$$
$$\times(1+\operatorname{erf}A_1)\}\left.\frac{\partial^2 f_2}{\partial R^2}\right|_{R=1}-4\exp(-A_1^2)\left.\frac{\partial^2 f_1}{\partial Z^2}\right|_{R=1}$$
$$+4\exp(-A_1^2)\left.\frac{\partial^2 f_2}{\partial Z^2}\right|_{R=1}, \tag{27}$$

$$A'(Z)=\mathrm{d}A_1/\mathrm{d}Z,$$
$$\{2A_1\exp(-A_1^2)+\pi^{1/2}(1+2A_1^2)(1+\operatorname{erf}A_1)\}\left.\frac{\partial^2 f_2}{\partial R^2}\right|_{R=1}$$
$$=\{2A_1\exp(-A_1^2)-\pi^{1/2}(1+2A_1^2)\operatorname{erfc}A_1\}\left.\frac{\partial^2 f_1}{\partial R^2}\right|_{R=1}$$
$$+2A_3\exp(-A_1^2)\{f_1(1,Z)-f_2(1,Z)\}$$
$$+2\{4A_1\exp(-A_1^2)-\pi^{1/2}(1+2A_2+4A_1^2)\operatorname{erfc}A_1\}\left.\frac{\partial f_1}{\partial R}\right|_{R=1}$$
$$+3\pi^{1/2}A_1^2+12\pi^{1/2}A_2+\{4A_1\exp(A_1^2)$$
$$-\pi^{1/2}(8A_2+4A_1^2+1/4)\operatorname{erfc}A_1\}f_1(1,Z)$$
$$-\{6A_1\exp(-A_1^2)+\pi^{1/2}(1+4A_2+6A_1^2)$$
$$\times(1+\operatorname{erf}A_1)\}\left.\frac{\partial f_2}{\partial R}\right|_{R=1}-3\{A_1\exp(-A_1^2)$$
$$+\pi^{1/2}(A_1^2+4A_2)(1+\operatorname{erf}A_1)\}f_2(1,Z)/2$$
$$+8A_1A_2(1-A_2)\exp(-A_1^2)\{f_1(1,Z)-f_2(1,Z)\}$$
$$+\pi^{1/2}\operatorname{erfc}(A_1)\left.\frac{\partial^2 f_1}{\partial Z^2}\right|_{R=1}$$
$$-\pi^{1/2}(1+\operatorname{erf}A_1)\left.\frac{\partial^2 f_2}{\partial Z^2}\right|_{R=1}. \tag{28}$$

In principle, other coefficients A_4, A_5 etc can also be determined, but the algebra becomes extremely complicated. Along with the unknowns in the moving boundary, the unknowns in the temperature solution in (10) are also determined.

We replace $I_0(x)$ in (10) by the first two terms of the asymptotic series in (12), use series expansions for f_1 and f_2 given in (18) and (19) and integrate term by term. $T(R, Z, V)$ is given by

$$\begin{aligned}
T(R, Z, V) = &-\tfrac{1}{2}\left[R^{1/2}\left\{f_1(1, Z)+(R-1)\left.\frac{\partial f_1}{\partial R}\right|_{R=1}\right.\right.\\
&\left.+\frac{(R-1)^2}{2}\left.\frac{\partial^2 f_1}{\partial R^2}\right|_{R=1}\right\}+\frac{V^2 f_1(1, Z)}{16R^{3/2}}+\frac{R^{1/2}V^2}{4}\left.\left.\frac{\partial^2 f_1}{\partial Z^2}\right|_{R=1}\right]\\
&\times\left[\operatorname{erf}\frac{(R-1)}{V}-1\right]+\frac{1}{2\pi^{1/2}}\left[-VR^{1/2}\left.\frac{\partial f_1}{\partial R}\right|_{R=1}\right.\\
&-(R-1)\,VR^{1/2}\left.\frac{\partial^2 f_1}{\partial R^2}\right|_{R=1}-VR^{-1/2}f_1(1, Z)-VR^{-1/2}\\
&\times(R-1)\left.\left.\frac{\partial f_1}{\partial R}\right|_{R=1}\right]\exp\left\{-\frac{(R-1)^2}{V^2}\right\}+\left(\frac{R^{1/2}V^2}{2\pi^{1/2}}\left.\frac{\partial^2 f_1}{\partial R^2}\right|_{R=1}\right.\\
&\left.+R^{-1/2}V^2\left.\frac{\partial f_1}{\partial R}\right|_{R=1}\right)\left[-\frac{(R-1)}{2V}\exp\left\{-\frac{(R-1)^2}{V^2}\right\}\right.\\
&\left.+\frac{\pi^{1/2}}{4}\operatorname{erfc}\left\{\frac{(R-1)}{V}\right\}\right]+\frac{1}{2}\left\{f_2(1, Z)+\frac{V^2}{16R^2}f_2(1, Z)\right.\\
&+\frac{(R-1)^2}{2}\left.\frac{\partial^2 f_2}{\partial R^2}\right|_{R=1}+(R-1)\left.\frac{\partial f_2}{\partial R}\right|_{R=1}+\frac{V^2}{4}\\
&\times\left.\left.\frac{\partial^2 f_2}{\partial Z^2}\right|_{R=1}\right\}\left[1+\operatorname{erf}\left\{\frac{(R-1)}{V}\right\}\right]\\
&-\frac{\pi^{1/2}}{2}\left\{-\frac{V}{2R}f_2(1, Z)-\frac{(R-1)V}{2R}\left.\frac{\partial f_2}{\partial R}\right|_{R=1}-V\left.\frac{\partial f_2}{\partial R}\right|_{R=1}\right.\\
&\left.-V(R-1)\left.\frac{\partial^2 f_2}{\partial R^2}\right|_{R=1}\right\}\exp\left\{-\frac{(R-1)^2}{V^2}\right\}\\
&-\frac{1}{4}\left\{-\frac{V^2}{8R^2}f_2(1, Z)+\frac{V^2}{2R}\left.\frac{\partial f_2}{\partial R}\right|_{R=1}+\frac{V^2}{2}\left.\frac{\partial^2 f_2}{\partial R^2}\right|_{R=1}\right\}\cdot\\
&\left[\frac{2(R-1)}{\pi^{1/2}V}\exp\left\{-\frac{(R-1)^2}{V^2}\right\}-\frac{\pi^{1/2}}{4}-\frac{\pi^{1/2}}{4}\operatorname{erf}\left\{\frac{(R-1)}{V}\right\}\right]\\
&+\text{terms of the form } (R-1)^m V^n, \text{ where } m+n>2.
\end{aligned}\tag{29}$$

The above temperature solution is valid for small values of V, $|R-1|$ and $|Z|$. In obtaining (29), whenever the limit of integration is R/V, it has been taken as infinity and this is justified as the integrals or error function integrals [1].

If only two terms of the asymptotic series in (12) are considered and $f_1(R, Z)$ is defined by (17) then the integrands in (13) and (14) do not have any singularity. If

more than two terms of (12) are to be considered then $f_1(R, Z)$ is to be suitably redefined such that the integrands in (13) and (14) do not have any singularity. However it may be noted that the asymptotic series in (12) does not converge.

4. Heat conduction without phase change

The short time temperature solution of pure heat conduction problem with $f(R, Z)$ as the initial temperature of the solid is given by (10) in which the unknown $f_2(R, Z)$ can be determined with the help of the boundary condition prescribed at $R=1$ which could be temperature prescribed or flux prescribed or radiation type. The temperature is still given by (29) with $f_2(R, Z)$ determined with the help of the prescribed boundary condition.

5. Some analytical checks

If $Q_0(Z)=0$ in (22) and the melting commences at $V=0$ then $f_1(1, Z)$ should be equal to unity for all Z. In this case, from (23), we get

$$f_1(1, Z)=f_2(1, Z)=1. \tag{30}$$

Equation (30) implies that $T(1, Z, 0)=1$ which it should be as this is the requirement of physics of the problem. It can be checked from (22) and (26) that (30) together with $Q_0(Z)=0$ implies that $A_1(Z)=0$ and $\partial f/\partial R|_{R=1}=\partial f_2/\partial R|_{R=1}$. For $V\ll 1$, $A_2(Z)$ will be the leading term in the moving boundary and it should be negative if melting has commenced at $V=0$. From (25) it can be seen that if

$$Q_1(Z)>\partial f/\partial R|_{R=1} \tag{31}$$

then $A_2(Z)$ will be negative. Equation (31) is the requirement of the physics of the problem for the starting of melting and it can be written down independently without doing any mathematical calculations. This provides a useful check on the method of solution and the algebra involved.

Some numerical checks will also be given below.

6. Numerical results and discussion

There does not seem to be any systematic way to determine analytically the radius of convergence of the series for the moving boundary as firstly, the series cannot be determined completely; secondly, even the rough estimates for the coefficients A_1, A_2, A_3 etc are difficult to obtain and thirdly the asymptotic expansion of $I_0(x)$ used earlier in determining these coefficients does not converge. In such a situation, a pertinent question can be asked: What is the range of time for which this short time solution is valid. The criterion for the validity of the range of time for which the

moving boundary solution is valid is simple. If the coefficients in the moving boundary are decreasing in absolute value (for fixed Z and given parameter values) then by calculating $|A_n(Z)\,V^n|$ (where A_n is the last term calculated in the moving boundary) it can be easily determined whether it makes any significant contribution to $\sum_{n=1}^{3} A_n(Z)V^n$ or not. If it does not make such a contribution then the moving boundary solution is valid atleast for this particular V. If the coefficients are not decreasing even then the solution is valid but the length of the time interval has to be extremely small. This criterion has been checked with the help of numerical schemes in [23], in which the analytical results for one-dimensional radially symmetric solidification problems were checked with the help of numerical schemes and were found in very good agreement. Numerical schemes for two-dimensional problems are fairly complicated and this infinite flux problem has not been tackled earlier by numerical schemes. It is hoped that the results presented in this paper can be used for starting and checking the numerical schemes.

Temperature solution given in (29) can be used for $V<1/2$ (as in the error function integrals the limit of integration R/V has been taken as ∞), $|R-1|$ small and $|Z|<1$.

In table 1 and figures 1 and 2, the data are as follows:

$$\lambda=0.6928,\ f_1(R,\ Z)=0.98-0.5(R-1)^2\exp(-Z^2),$$

$$Q(Z,\ V)=0.5\{1+\exp(-Z^2)\}\,\{1/V+0.5+0.25V\}.$$

In table 1, the isotherm condition (5) has been checked numerically. For given values of V and Z, a value of R can be obtained from the relation

$$S=1+\sum_{n=1}^{3} A_n(Z)V^n \tag{32}$$

and this value of R was substituted in (29) together with given values of V and Z. The resulting temperature values are given below.

It can be easily seen that the temperature is almost unity at the moving boundary (the error is significant). This checking of isotherm condition provides a good check on the method of solution and the algebra involved.

Table 1. Checking the isotherm condition.

Z/V	S	Temperature	$A_1(Z)$	$A_2(Z)$	$A_3(Z)$
0.0/0.05	0.9647	1.0008	−0.6909	−0.2997	−0.0532
0.0/0.10	0.9278	0.9992	−0.6909	−0.2997	−0.0532
0.6/0.05	0.9703	0.9997	−0.5843	−0.1652	−0.0101
0.6/0.10	0.9399	1.0007	−0.5843	−0.1652	−0.0101

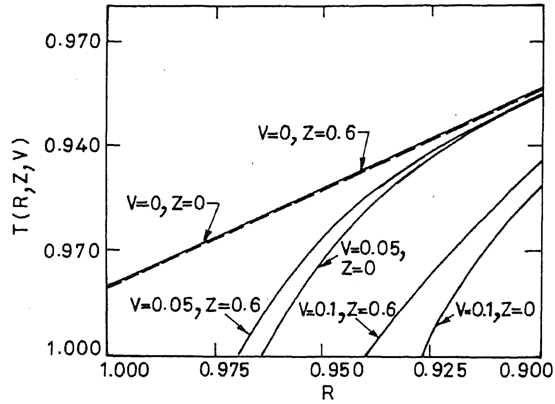


Figure 1. Temperature in the solid vs R for different values of V and Z. $\lambda = 0.6928$, $f_1(R, Z) = 0.98 - 0.5(R-1)^2 \exp(-Z^2)$, $Q(Z, V) = 0.5(1/V + 0.5 + 0.25V) \exp(-Z^2)$.

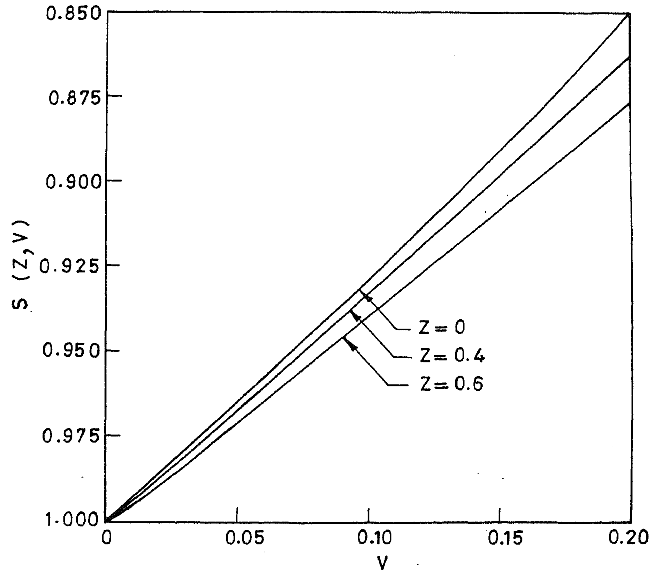


Figure 2. Melting front vs V for different values of Z. Data as in figure 1.

In figure 1, the temperature in the solid vs R has been plotted for different V and Z and this has been compared with the initial temperature of the solid. In figure 2 the melting front vs V has been plotted for $Z = 0$, $Z = 0.4$ and $Z = 0.6$. The trend of the graphs in figures 1 and 2 suggests that the solution is valid atleast for the values of Z, R and V reported in these figures. The melted portion of the solid is quite substantial for use of direct applications or checking the numerical schemes.

Acknowledgement

The author wishes to thank the referee for some useful suggestions.

References

[1] Abramowitz M and Stegun I A, *Handbook of mathematical functions*, Dover Publications, New York (1965)

[2] Bankoff S G, Heat conduction or diffusion with change of phase; in *Advances in chemical engineering*, Academic Press, New York (1964) 5, 75–150

[3] Boley B A, A method of heat conduction analysis of melting and solidification problems, *J. Math. Phys.* **40** (1961) 300–313

[4] Boley B A, in: Ockendon J R and Hodgkins W R, Moving boundary problems in heat flow and diffusion, Clarendon Press, Oxford (1975)

[5] Carslaw H S and Jaeger J C, Conduction of heat in solids, Clarendon Press, Oxford (1948)

[6] Crank J and Gupta R S, A method for solving moving boundary problems in heat flow using cubic splines or polynomials, *J. Inst. Math. Appl.* **10** (1972) 19–33

[7] Cross M and Voller V R, Accurate solutions of moving boundary problems using enthalpy method, *Int. J. Heat Mass Transfer* **24** (1981) 545–555

[8] Crowley A B, Numerical solution of Stefan problems, *Int. J. Heat Mass Transfer* **21** (1979) 215–219

[9] Furzeland R M, A comparative study of numerical methods for moving boundary problems, *J. Inst. Math. Appl.* **26** (1980) 411–429

[10] Goodman T R, The heat balance integral—further considerations and refinements, *ASME J. Heat Transfer* **83** (1961) 81–85

[11] Gupta S C, Two-dimensional heat conduction with phase change in a semi-infinite mold, *Int. J. Engg. Sci.* **19** (1981) 137–146

[12] Gupta S C and Lahiri A K, Heat conduction with a phase change in a cylindrical mould, *Int. J. Engg. Sci.* **17** (1979) 401–307

[13] Muehlbauer J C and Sunderland J E, Heat conduction with freezing or melting, *Appl. Mech. Rev.* **18** (1965) 951–959

[14] Murray W D and Landis F, Numerical and machine solutions of transient heat conduction problems involving melting or freezing, Part I—Method of analysis and sample solutions, *ASME J. Heat Transfer* **81** (1959) 106–112

[15] Ockendon J R and Hodgkins W R, Moving boundary problems in heat flow and diffusion, Clarendon Press, Oxford (1975)

[16] Ozisik M N and Uzzell J C (Jr.), Exact solution for freezing in cylindrical symmetry with extended temperature range, *ASME J. Heat Transfer* **101** (1979) 331–334

[17] Pedroso R I and Domoto G A, Perturbation solutions for spherical solidification of saturated liquids, *ASME J. Heat Transfer* **95** (1973) 42–46

[18] Ruddle R W, The solidification of castings, Institute of Metals, London (1957)

[19] Schulze H J, Beckett P M, Howarth J A and Poots G, Analytical and numerical solutions to two-dimensional moving interface problems with applications to the solidification of killed steel ingots, *Proc. Roy. Soc.* **A383** (1983) 312–343

[20] Sikarskie D L and Boley B A, The solution of a class of two-dimensional melting and solidification problems, *Int. J. Solids Struct.* **1** (1965) 207–234

[21] Tao L N, On solidification problems including the density jump at the moving boundary, *Q. J. Mech. Appl. Math.* **32** (1979) 175–185

[22] Tiller W A, Grain size control during ingot solidification, *Trans. Metall. Soc. A.I.M.E.* **224** (1962) 448–459

[23] Vellingiri R, Solutions of some radially symmetric problems of heat conduction with phase change, Ph.D. dissertation, Indian Institute of Science, Bangalore (1983)

Proc. Indian Acad. Sci. (Math. Sci.), Vol. 95, No. 1, September 1986, pp. 13–21.

Instability of a planar liquid layer in an alternating longitudinal magnetic field with non-zero mean

S P BHATTACHARYYA and M ABBAS*

Department of Mathematics, * Department of Computer Science and Engineering, Indian Institute of Technology, Powai, Bombay 400 076, India

MS received 25 March 1985; revised 30 March 1986

Abstract. The conducting liquid interface is found to undulate in an alternating magnetic field. It was shown earlier that if $M = |B_0|^2/\mu\eta\omega$, B_0, ω, μ and η being the amplitude (complex) of the alternating longitudinal magnetic field imposed at the interface, the angular frequency of the field, the magnetic permeability and the viscosity respectively, and if M_c was the critical value of M then the planar layer was stable or unstable according as $M < M_c$ or $M > M_c$. In this paper we have determined the stability criterion when in addition to the alternating longitudinal field there acts a uniform field in the same direction. After comparing our results with those obtained earlier, in the absence of the uniform field, we find that the additional uniform field has a significant destabilizing effect.

Keywords. Stability; alternating magnetic field; interface; geomagnetic field.

1. Introduction

In certain cases, where magnetic fields are used for levitation, melting of metals and shaping of the interface to form materials like glass, it has been found that the alternating field causes undulations [3, 8, 4, 5]. Such undulation takes place even in the case of hydromagnetic surface waves [7].

In order to determine the possible causes of the undulations, an analysis based on a 'lumped parameter' model [4] taking thermal effects into account and also a linear stability analysis based on usual MHD model neglecting thermal effects [5] was carried out. It was suggested that electromechanical motion driven by eddy current heating and the cooling from the upper surface could cause such undulations. The effect of such additional uniform field such as the geomagnetic field, acting in the layer along with the alternating field has not been studied earlier. We have recently investigated the stability of a conducting planar liquid layer in a longitudinal alternating magnetic field imposed at the interface with a uniform field acting in the transverse direction [2]. It is found that such an additional uniform field has considerable destabilizing effect. In order to make a comparative study, we have investigated the stability when the uniform magnetic field acts in the same direction as the longitudinal alternating field. In this paper, we have presented in §2 the linear stability analysis and the numerical results.

2. Linear stability analysis and the numerical results

We take a rectangular cartesian system of axes $oxyz$ with o as the origin on the interface, ox axis along the outward normal to the layer and oz axis in the direction of the imposed longitudinal field real $B_0 \exp(j\omega t)$, where B_0 is the amplitude (complex) of the field, ω the angular frequency, t the time and $j=\sqrt{-1}$. We assume the uniform field B_1 to be acting in the liquid in the same direction so that if $\delta=(2/\omega\mu\sigma)^{1/2}$, where σ is the conductivity, then the magnetic field $\mathbf{B}$ in the liquid in the absence of any motion is given by

$$\begin{aligned} \mathbf{B} &= \mathbf{B}_e + \mathbf{B}_1, \ \mathbf{B}_1 = B_1 \mathbf{i}_z \ (B_1 \text{ real}), \\ \mathbf{B}_e &= B_0 \exp(1+j)\frac{x}{\delta} \exp(j\omega t)\mathbf{i}_z . \end{aligned} \tag{1}$$

The unit vectors in the direction of x, y and z axes are being denoted by $\mathbf{i}_x$, $\mathbf{j}_y$ and $\mathbf{i}_z$ respectively.

In the static state, we take p_0 as the hydromagnetic pressure, whereas in the disturbed state we take p_0+p as the pressure, $\mathbf{v}$ as the velocity and $\mathbf{B}+\mathbf{b}$ as the magnetic field where p, $\mathbf{v}$ and $\mathbf{b}$ are assumed to be small in magnitude. The linearized MHD equations are

$$\begin{aligned} \rho \frac{\partial \mathbf{v}}{\partial t} = -\nabla p + \eta \nabla^2 \mathbf{v} + \frac{1}{\mu}(\mathbf{B}_e + \mathbf{B}_1)\cdot\nabla\mathbf{b} \\ + \frac{1}{\mu}(\mathbf{b}\,.\,\nabla)\mathbf{B}_e, \end{aligned} \tag{2}$$

$$\frac{\partial \mathbf{b}}{\partial t} = \frac{1}{\mu\sigma}\nabla^2\mathbf{b} + (\mathbf{B}_0+\mathbf{B}_1)\cdot\nabla\mathbf{v} - (\mathbf{v}\cdot\nabla)\mathbf{B}_e \tag{3}$$

$$\nabla\cdot\mathbf{v} = \nabla\cdot\mathbf{b} = 0, \tag{4}$$

where ρ and η are density and viscosity respectively.

We seek the solution of (2)–(4) in the following approximate forms

$$v = \text{real } v(x)\exp(st - jk_y y - jk_z z), \tag{5}$$

$$p = \text{real } p(x)\exp(st - jk_y y - jk_z z), \tag{6}$$

$$\begin{aligned} b = \text{real}[&b^+(x)\exp(st + j\omega t - jk_y y - jk_z z), \\ &+ b^-(x)\exp(st - j\omega t - jk_y y - jk_z z), \\ &+ c(x)\exp(st - jk_y y - jk_z z)]. \end{aligned} \tag{7}$$

Here s has been taken as complex. It may be noted that due to the presence of $\mathbf{B}_e$ as given by (1) in (2) and (3), the solution can be expressed as a series of terms, $\exp(st-jk_y y-jk_z z)$, $\exp(st \pm j\omega t - jk_y y - jk_z z)$, $\exp(st \pm 2j\omega t - jk_y y - jk_z z)$ etc. Assuming the viscous damping and inertia to inhibit the growth of terms with frequencies higher than that given by $\text{Im}(s)$ ($\text{Im}(s)$ being the imaginary part of s).

We have retained terms involving only $\exp(st - jk_y y - jk_z z)$ in (5) and (6) as approximate solutions.

We observe from physical considerations that since the velocity field and the magnetic field in a conducting fluid cannot be considered to be decoupled, the terms $b^{\pm}(x)\exp(st \pm j\omega t - jk_y y - jk_z z)$ and also $c(x)\exp(st - jk_y y - jk_z z)$ in (7) must be retained to balance the equations (2) and (3) to the lowest order.

Here we present certain considerations similar to those presented by McHale and Melcher [5] in the introduction and followed in the subsequent calculations. We assume $\omega \simeq 2$ kHz and $|\mathrm{Im}(s)|$ to be small due to viscosity. As $|\mathrm{real}(s)|$ may be assumed to be sufficiently small if we are interested in the study of the state when it is close to the marginal state, we can take $|s| \ll \omega$ in the following analysis.

We substitute (5), (6), (7) in (2), (3), (4) eliminating p and introducing the following non-dimensional variables denoted by dashes

$$(x, y, z) \equiv \delta(x', y', z), \quad (k, k_z) \equiv \delta^{-1}(k', k_z'),$$

$$(v_x, v_y, v_z) \equiv (\mu\sigma\delta)^{-1}(v_x', v_y', v_z'),$$

$$(b_x^+, b_y^+, b_z^+) \equiv B_0(b_x^{+\prime}, b_y^{+\prime}, b_z^{+\prime}),$$

$$(b_x^-, b_y^-, b_z^-) \equiv B_0(b_x^{-\prime}, b_y^{-\prime}, b_z^{-\prime}),$$

$$(c_x, c_y, c_z) \equiv |B_0|(c_x', c_y', c_z'),$$

also the following non-dimensional parameters

$$P_m = \mu\sigma\eta/2\rho, \quad S' = S/\omega P_m, \quad B_1' = B_1/|B_0|, \quad M = |B_0|^2/\mu\eta\omega$$

and obtain the following equations in non-dimensional forms after neglecting $S'P_m$ due to the considerations stated above. We have

$$\begin{aligned}&[(D^2 - k^2)(D^2 - k^2 - s) + Mk_z^2(2B_1^2 + \exp(2x))]v_x\\&\quad + 4Mk_z\{b_x^+\exp((1-j)x) - b_x^-\exp((1+j)x)\} = 0,\end{aligned} \tag{8}$$

$$(D^2 - k^2 - 2j)b_x^+ = \tfrac{1}{2}(jk_z)\exp((1+j)x)v_x, \tag{9}$$

$$(D^2 - k^2 + 2j)b_x^- = \tfrac{1}{2}(jk_z)\exp((1-j)x)v_x, \tag{10}$$

$$(D^2 - k^2)c_x = jk_z B_1 v_x, \tag{11}$$

$$D = \mathrm{d}/\mathrm{d}x.$$

In the above equations dashes have been dropped for simplicity.

We assume the layer to be thick and the liquid air interface rigid and perfectly conducting. We take the boundary conditions as

$$v_x = Dv_x = b_x^+ = b_x^- = c_x = 0 \quad \text{at } x = 0 \tag{12}$$

together with the conditions: v_x, Dv_x etc. tend to zero as $x \to -\infty$.

3. Validity of the principle of exchange of stability

After multiplying (8) by v_x^*, the complex conjugate of v_x and integrating with respect to x from $-\infty$ to 0, we obtain, after using (9) and (10)

$$\int_{-\infty}^{0} [|D^2v_x|^2 + (2k^2+s)|Dv_x|^2 + k^2(k^2+s)|v_x|^2$$
$$+ Mk_z^2\{2B_1^2 + \exp(2x)\}|v_x|^2]\mathrm{d}x,$$
$$-8Mj \int_{-\infty}^{0} [|Db_x^+|^2 - |Db_x^-|^2 + k^2\{|b_x^+|^2 - |b_x^-|^2\}$$
$$-2j\{|b_x^+|^2 + |b_x^-|^2\}]\mathrm{d}x = 0. \tag{13}$$

From (9) and (10) we get

$$\left.\begin{aligned}(D^2-k^2-2j)b_x^+ &= \frac{jk_z}{2}\exp((1+j)x)v_x\\ (D^2-k^2-2j)(b_x^-)^* &= -\frac{jk_z}{2}\exp((1+j)x)v_x^*\end{aligned}\right\}, \tag{14}$$

where $(b_x^-)^*$ is the complex conjugate of b_x^-. We find from (5) that $v(x)$ must be either real or purely imaginary, since we are interested in finding the criterion under which a single mode of perturbation grows or decays. Hence from (14), we get $b_x^- = -(b_x^+)^*$ when v_x is real and $b_x^- = (b_x^+)^*$ when v_x is purely imaginary. In both cases we have $|Db_x^+| = |Db_x^-|$ and $|b_x^+| = |b_x^-|$ so that (13) finally reduces to

$$\int_{-\infty}^{0} [|D^2v_x|^2 + (2k^2+s)|Dv_x|^2 + k^2(k^2+s)|v_x|^2$$
$$+ Mk_z^2\, 2B_1^2 + \exp(2x)\,|v_x|^2]\mathrm{d}x,$$
$$-32M \int_{-\infty}^{0} |b_x^+|^2\mathrm{d}x = 0. \tag{15}$$

If we take the imaginary part of the above equation, we get

$$\mathrm{Im}(s)\int_{-\infty}^{0} \{|Dv_x|^2 + k^2|v_x|^2\}\mathrm{d}x = 0 \tag{16}$$

and hence $\mathrm{Im}(s) = 0$ for any mode of disturbance. Thus the principle of exchange of stabilities is valid indicating that the marginal state must be convective.

To solve equations (8)–(11) subject to the boundary condition (12) with the physical consideration that all disturbances should remain bounded in magnitude at an infinitely large distance away from the surface, we follow a method presented in [1] and suitably extended in [5]. We observe that if we set $\xi = \exp(2x)$ in (8)–(11) the differential equations can be transformed into equations for which $\xi = 0$ appears to be a regular singularity. We seek solutions of (8)–(11) in series of

powers of $\xi = \exp(2x)$ and write formally

$$V_x = A_0 \exp(r_n x) \sum_{m=0}^{\infty} D_m \exp(2mx), \tag{17a}$$

$$b_x^+ = A_0 \exp\{(r_n + 1 + j)x\} \sum_{m=0}^{\infty} E_m^+ \exp(2mx), \tag{17b}$$

$$b_x^- = A_0 \exp\{(r_n + 1 - j)x\} \sum_{m=0}^{\infty} E_m^- \exp(2mx), \tag{17c}$$

$$c_x = A_0 \exp(r_n x) \sum_{m=0}^{\infty} F_m \exp(2mx) \tag{17d}$$

Since the set of equations (8)–(11) constitutes a differential equation of order ten, the success of the method lies in determining ten linearly independent solutions of (8)–(11), which should correspond to ten roots of r_n all of which need not be distinct.

Substituting (17a–d) in (8)–(11) and equating the coefficients of the same powers of $\exp(2x)$ we have the following set of equations

$$\{(r_n^2 - k^2)(r_n^2 - k^2 - s) + 2B_1^2 M k_z^2\} D_0 = 0, \tag{18a}$$

$$[\{r_n + (1 + j)\}^2 - k^2 - 2j] E_0^+ = \tfrac{1}{2}(jk_2) D_0, \tag{18b}$$

$$[\{r_n + (1 - j)\}^2 - k^2 + 2j] E_0^- = \tfrac{1}{2}(jk_z) D_0, \tag{18c}$$

$$(r_n^2 - k^2) F_0 = jk_z B_1 D_0, \tag{18d}$$

and for $m \geq 1$

$$[\{(r_n + 2m)^2 - k^2\}\{(r_n + 2m)^2 - k^2 - s\} + 2B_1^2 M k_z^2] D_m$$
$$+ M k_z^2 D_{m-1} + 4Mk_z(E_{m-1}^+ - E_{m-1}^-) = 0, \tag{19a}$$

$$[\{r_n + (1 + j) + 2m\}^2 - k^2 - 2j] E_m^+ = \tfrac{1}{2}(jk_z) D_m, \tag{19b}$$

$$[\{r_n + (1 - j) + 2m\}^2 - k^2 + 2j] E_m^- = \tfrac{1}{2}(jk_z) D_m, \tag{19c}$$

$$[(r_n + 2m)^2 - k^2] F_m = jk_z B_1 D_m. \tag{19d}$$

From (18a–d) we find that after eliminating D_0, E_0^+, D_0^- and F_0 we get an algebraic equation of degree ten in r_n. Hence for an arbitrary choice of D_0, E_0^+, E_0^- and F_0 we can calculate D_m, E_m^+, E_m^- and F_m for all $m \geq 1$ after taking one of the roots of r_n, giving finally one particular solution. In this way all the ten linearly independent solutions can be obtained though there may arise a difficulty in case of two or more roots of r_n happen to be equal in which case we may have to use the method of Frobenius with necessary modifications.

We can simplify the above procedure if we consider that our solution must satisfy the condition of boundedness as $x \to -\infty$ and therefore we need consider only those particular solutions for which real $(r_n) > 0$. Moreover we find that instead of

considering a single set of values of D_0, E_0^+, E_0^- and F_0 we can consider different sets of values of D_0, E_0^+, E_0^- and F_0 to obtain particular solutions, provided such solutions are linearly independent. We therefore present the following scheme.

Solutions 1, 2.

$$D_0 = 1;\ (r_n^2 - k^2)(r_n^2 - k^2 - s) \times 2B_1^2 M k_z^2 = 0,$$

$$r_n = r_1,\ r_2,\ \text{real}(r_1, r_2) > 0.$$

Solution 3.

$$D_0 = E_0^- = F_0 = 0;\ E_0^+ = 1,\ (r_n + 1 + j)^2 = k^2 + 2j,$$

$$r_n = r_3;\ \text{real}(r_3) > 0.$$

Solution 4.

$$D_0 = E_0^+ = F_0 = 0;\ E_0^- = 1,\ (r_n + 1 - j)^2 = k^2 - 2j,$$

$$r_n = r_4;\ \text{real}(r_4) > 0.$$

Solution 5.

$$D_0 = E_0^+ = E_0^- = 0;\ F_0 = 1;\ r_n^2 - k^2 = 0,$$

$$r_n = r_5;\ \text{real}(r_5) > 0.$$

In case we get five distinct values of r_n we find that we can express the general solution as

$$\begin{aligned}
V_x = {} & A_1 \exp(r_1 x)[1 + D_1^{(1)} \exp(2x) + D_2^{(1)} \exp(4x) + \ldots] \\
& + A_2 \exp(r_2 x)[1 + D_1^{(2)} \exp(2x) + D_2^{(2)} \exp(4x) + \ldots] \\
& + A_3 \exp(r_3 x)[D_1^{(3)} \exp(2x) + D_2^{(3)} \exp(4x) + \ldots] \\
& + A_4 \exp(r_4 x)[D_1^{(4)} \exp(2x) + D_2^{(4)} \exp(4x) + \ldots] \\
& + A_5 \exp(r_5 x)[D_1^{(5)} \exp(2x) + D_2^{(5)} \exp(4x) + \ldots]
\end{aligned}$$

$$\begin{aligned}
b_x^+ = {} & A_1 \exp\{(r_1 + 1 + j)x\}[E_0^{+(1)} + E_1^{+(1)} \exp(2x) \\
& + E_2^{+(1)} \exp(4x) + \ldots] \\
& + A_2 \exp\{(r_2 + 1 + j)x\}[E_0^{+(2)} + E_1^{+(2)} \exp(2x) \\
& + E_2^{+(2)} \exp(4x) + \ldots] \\
& + A_3 \exp\{(r_3 + 1 + j)x\}[1 + E_1^{+(3)} \exp(2x) + E_2^{+(3)} \exp(4x) + \ldots] \\
& + A_4 \exp[(r_4 + 1 + j)x][E_1^{+(4)} + \exp(2x) + E_2^{+(4)} \exp(4x) + \ldots] \\
& + A_5 \exp[(r_5 + 1 + j)x][E_1^{+(5)} + \exp(2x) + E_2^{+(5)} \exp(4x) + \ldots]
\end{aligned}$$

and similarly b_x^- and c_x can be expressed as series where all the quantities $D_1^{(1)}$, $D_2^{(1)}$ etc are obtained from (18a–d) and the recurring relations (19a–d).

Using the boundary condition (12) after considering the solutions in the above forms and eliminating A_1, A_2 etc, we get a determinantal equation denoted by

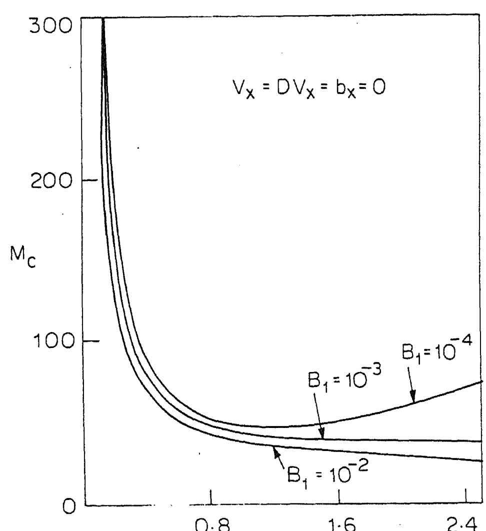
$$Q(M, B_1, s, k, k_z) = 0, \tag{20}$$

where Q is a 5×5 determinant which reduces to a 4×4 determinant due to the particular form of the solution 5 when $D_m = E_m^+ = E_m^- = F_m = 0$ for $m \geq 1$. It may be pointed out that the same 4×4 determinant could have been obtained if we had confined to only equations (8)–(10). In the above 4×4 determinant every element is an infinite series of terms dependent upon the variables indicated in (20).

The critical value M_c of M for a particular value of k is given by the minimum of all the roots of M in (20) for all values of k_z with k kept fixed when $s = 0$. Since equations (8)–(11) can be transformed into the forms in which k_z appears only along with Mk_z^2 and not independently, to obtain the minimum value of M we set $k_z = k$ in (20).

We have calculated the values of M_c for $0 < k \leq 2.4$ then $B_1 = 10^{-2}$, 10^{-3} and 10^{-4}. The plots of M_c vs k have been presented in figure 1.

For smaller values of $k\,(k < 0.4)$ the changes in the roots of M_c in (20) are large for changes in the values of k and the roots of M_c have been located after fixing up M and observing the changes in the signs of the real and imaginary parts of Q as given in (20) for successive values of k from a set of values of k. The accuracy has been improved by taking large number of terms in each element of the determinant Q. For larger values of $k\,(k > 0.4)$ the changes in the roots of M_c in (20) are found to be small. To locate the roots of M_c we have repeated the above procedure by



fixing up k and considering the changes in signs of the real and imaginary parts of Q for successive values of M from a set of values of M.

4. Discussion

From figure 1 we find that for $B_1 = 10^{-2}$, 10^{-3} and 10^{-4} the values of M_c remain nearly the same for $k < 0.4$ (approximately) whereas for larger values of k there is a slow increase in the values of M_c for $B_1 = 10^{-4}$ and slow decrease in the values of M_c for $B_1 = 10^{-3}$ and 10^{-2} and hence the additional uniform field has destabilizing effect.

It may be concluded that the observed undulation of the liquid metal air interface in an alternating longitudinal magnetic field may be partly caused by the geomagnetic field or any other stray field in the laboratory. In particular, for a liquid thick layer of mercury at 20°C when the kinematic viscosity, the magnetic permeability and the density are taken as 10^{-7} m^2/s, $4\pi\, 10^{-7}$ henry/m and 10^7 gm/m^3 respectively and Earth's magnetic field is taken as 0.7 Gauss [6], we find that for a longitudinal alternating field with amplitude 1.5 Gauss and angular frequency 5 kHz, the value of M is obtained as 352 (approximately) whereas the value of $B_1 = 0.46$ (approximately). Thus according to the results presented by the plots of M_c vs k we find the layer to be unstable since as B_1 increases from 10^{-4} the critical value of M decreases and is less than 40. But according to the results obtained by McHale and Melcher [5], the layer is supposed to remain stable.

Acknowledgements

The numerical computation reported in this paper has been performed on CYBER 730 System (Control Data Corporation) installed at TIFR, Bombay. The investigation has been carried out under a scheme sponsored by the Department of Atomic Energy. We would like to thank the referee for valuable suggestions for improvement of the paper.

References

[1] Betchov R and Criminale O, *Stability of parallel flows*, Academic Press, New York (1967) Chapt. 3

[2] Bhattacharyya S P and Abbas M, Instability of a planar liquid layer in a magnetic field with alternating longitudinal and uniform transverse components, *Int. J. Engg. Sci.* **24** (1986) 779–783

[3] Fraser M E, Lu W K, Hamiclec A E and Murarka R, Surface tension measurements on pure liquid iron and nickel by an oscillating drop technique, *Metall. Trans.* **2** (1971) 817–823

[4] McHale E J and Melcher J R, Hydromagnetic instability of liquid metals in AC magnetic fields and augmentation of heat transfer, *Electr. Mech. Electromech.* **3** (1978) 197

[5] McHale E J and Melcher J R, Instability of a planar liquid layer in an alternating magnetic field, *J. Fluid Mech.* **114** (1982) 27–40

[6] Roberts P H, *An introduction to magnetohydrodynamics*, pp. 5 and 257, Longmans London (1967)
[7] Schaffer M, Hydromagnetic surface waves with alternating magnetic fields, *J. Fluid Mech.* **33** (1968) 337–351
[8] Sunderlam T, Thermal diffusion in kinetic and equilibrium measurements by levitation technique, *Metall. Trans.* **4** (1973) 575–583

Proc. Indian Acad. Sci. (Math. Sci.), Vol. 95, No. 1, September 1986, pp. 23–30.

Convective instability of a rotating layer of ferromagnetic fluid

S P BHATTACHARYYA and M ABBAS*

Department of Mathematics, *Department of Computer Science and Engineering, Indian Institute of Technology, Powai, Bombay 400 076, India

MS received 21 February 1986

Abstract. The instability of a hot horizontal layer of ferromagnetic fluid rotating about a vertical axis has been investigated when the Prandtl number $P<1$. Earlier it was shown that for $P>1$ the overstability cannot occur. In this paper the convective and overstable marginal states have been investigated separately for $P<1$ and it is found that though convective marginal state is possible for all a, the non-dimensional wave number, and N the Taylor number, the overstability is possible only if $N>(1+P)\pi^4/(1-P)$ and in case the condition is satisfied, overstability is possible for all those values of a which satisfy $a^2<[N(1-P)\pi^2/(1+P)]^{1/3}-\pi^2$. If $R_c^{(\mathrm{con})}$ and $R_c^{(\mathrm{o.s})}$ are the critical values of the convective and the overstable marginal states respectively, then it is also found that $R_c^{(\mathrm{con})} < R_c^{(\mathrm{o.s})}$ provided N is not sufficiently large.

Keywords. Stability; convective marginal state; overstability; ferromagnetic fluid.

1. Introduction

Ferromagnetic fluids are formed by suspending submicron sized magnetic particles in a medium like kerosene or water with the addition of some anticoagulating liquid like Oleic acid. It is found that the continuum theory is applicable to study the flow behaviour of such a fluid. Due to certain important applications of the ferromagnetic fluid in energy conversion devices etc., there is a great deal of interest in studying the flow characteristics of the fluid in recent years.

Neuringer and Rosenweig [3] investigated the heat transfer processes in ferromagnetic fluid. Finlayson [2] investigated the stability of a hot layer of ferromagnetic fluid and showed that the principle of exchange of stabilities was valid in case the boundaries were free. Das Gupta and Gupta [1] studied the instability of a hot layer of ferromagnetic fluid rotating about a vertical axis and showed that the overstability could not occur if the Prandtl number $P > 1$. Since it appeared that there was the possibility of having overstability if $P < 1$, we have investigated in this paper the conditions under which the critical marginal state is convective or overstable when $P < 1$. To shorten our discussion, we have referred to the non-dimensional equations incorporating the normal mode analysis and the boundary conditions for free boundaries as were used by Das Gupta and Gupta [1] in §2. The critical values of the Rayleigh numbers $R_c^{(\mathrm{con})}$ and $R_c^{(\mathrm{o.s})}$ for the convective and the overstable marginal states have been calculated separately in

§3. In §4, we have presented the numerical results in two tables, and the plots $R_c^{(\text{con})}$ and $R_c^{(\text{o.s})}$ vs a^2 followed by a discussion.

2. Nondimensional equations and the boundary conditions

We consider a horizontal layer of incompressible ferromagnetic fluid heated from below and rotating about a vertical axis. Following the linearization of the basic equations, non-dimensionalization and using the normal mode analysis as were done by Das Gupta and Gupta [3], we present

$$\frac{\partial}{\partial t}(D^2-a^2)W = -N^{1/2}D\zeta + aR^{1/2}[M_1 D\phi - (1+M_1)T] + (D^2-a^2)^2 W, \tag{1}$$

$$\frac{\partial \zeta}{\partial t} = (D^2-a^2)\zeta + N^{1/2}DW, \tag{2}$$

$$P\frac{\partial T}{\partial t} = (D^2-a^2)T + aR^{1/2}W, \tag{3}$$

$$D^2\phi - a^2 M_3\phi - DT = 0, \tag{4}$$

and the boundary conditions

$$W = D^2W = T = D\phi = 0 \quad \text{at} \quad z = \pm 1/2. \tag{5}$$

We have used the same notations as used by Das Gupta and Gupta [1] and we have taken N to represent the Taylor number $4\Omega^2 d^4/r^2$. For simplification, we have dropped M_2 in the above equations as $M_2 \simeq 10^{-6}$ ([2] and [1]).

Since the solutions of (1)–(5) can be separated into even and odd modes, we take them as

$$W = A\exp(\sigma_n t)\,{}^{\cos}_{\sin}\, r_n z, \quad T = B\exp(\sigma_n t)\,{}^{\cos}_{\sin}\, r_n z,$$

$$\phi = \frac{C}{r_n}\exp(\sigma_n t)\,{}^{\sin}_{\cos}\, r_n z, \quad \zeta = E\exp(\sigma_n t)\,{}^{\sin}_{\cos}\, r_n z, \tag{6}$$

where the upper terms signify even mode for which $r_n = (2n-1)\pi$ ($n = 1, 2, 3, \ldots$) and the lower terms, for which $r_n = 2n\pi$ ($n = 1, 2, 3, \ldots$), signify odd mode of the perturbation.

Substituting (6) into (1)–(4) and eliminating A, B, C and E, we get two characteristic equations for even and odd modes, which we combine into one, as

$$U\sigma_n^3 + V\sigma_n^2 + W\sigma_n + X = 0, \tag{7}$$

where

$$U = P(n^2\pi^2 + a^2)\ (n^2\pi^2 + a^2 M_3), \tag{8}$$

$$V = (2P+1)\ (n^2\pi^2 + a^2)^2\ (n^2\pi^2 + a^2M_3), \tag{9}$$

$$W = (P+2)\ (n^2\pi^2 + a^2)^3\ (n^2\pi^2 + a^2M_3) + NPn^2\pi^2(n^2\pi^2 + a^2M_3) - a^2R[n^2\pi^2 + a^2M_3(1+M_1)], \tag{10}$$

$$X = (n^2\pi^2 + a^2)\ (n^2\pi^2 + a^2M_3) + Nn^2\pi^2(n^2\pi^2 + a^2)\ (n^2\pi^2 + a^2M_3) - a^2R(n^2\pi^2 + a^2)\ [n^2\pi^2 + a^2M_3(1+M_1)], \tag{11}$$

where n is a positive integer, even in the case of an odd mode and odd in the case of an even mode of perturbation.

We set $\sigma_n = \sigma_n^{(r)} + i\sigma_n^{(i)}$; $\sigma_n^{(r)}$ and $\sigma_n^{(i)}$ being the real and the imaginary parts of σ_n and obtain from (7), after separating the real and imaginary parts

$$U\sigma_n^{(r)}\ [\{\sigma_n^{(r)}\}^2 - 3\{\sigma_n^{(i)}\}^2] + V[\{\sigma_n^{(r)}\}^2 - \{\sigma_n^{(i)}\}^2] + W\sigma_n^{(r)} + X = 0 \tag{12}$$

and

$$U\sigma_n^{(i)}\ [3\{\sigma_n^{(r)}\}^2 - \{\sigma_n^{(i)}\}^2] + \sigma_n^{(i)}[2V\sigma_n^{(r)} + W] = 0 \tag{13}$$

If the marginal state is convective, we obtain from (12) after setting $\sigma_n^{(i)} = 0$, as

$$U\{\sigma_n^{(r)}\}^3 + V\{\sigma_n^{(r)}\}^2 + W\sigma_n^{(r)} + X = 0. \tag{14}$$

We note that in the convective marginal state when at least one root of (14) is zero, other real roots must be non-positive. Thus, in the case of the convective marginal state, we must have

$$X = 0, \quad W \geq 0 \tag{15}$$

as $U \geq 0$, and $V \geq 0$.

Again for overstable marginal state, after eliminating $\sigma_n^{(i)}$ from (12) and (13), under the assumption $\sigma_n^{(i)} \neq 0$, we have

$$8U^2\{\sigma_n^{(r)}\}^3 + 8UV\{\sigma_n^{(r)}\}^2 + 2(WU + V^2)\ \{\sigma_n^{(r)}\} + \{WV - UX\} = 0. \tag{16}$$

In the case of the overstability when at least one of the roots of $\sigma_n^{(r)}$ in (16) is zero, other real roots must be non-positive and moreover at least one $\sigma_n^{(i)}$ obtained from (13) must be real. Noting that U and V are positive, we obtain from (16) and (13) that in the overstable marginal state

$$WV - UX = 0, \quad WU + V^2 \geq 0 \tag{17}$$

and

$$(\sigma_n^{(i)})^2 = W/U > 0. \tag{18}$$

2.1 Evaluation of $R_c^{(\mathrm{con})}$

We use the conditions (15) to get the value of $R_c^{(\mathrm{con})}$. Taking $X = 0$, we have, after setting $R = R_n^{(\mathrm{con})}$,

$$R_n^{(\mathrm{con})} = \frac{[(n^2\pi^2 + a^2)^3 + Nn^2\pi^2]\ (n^2\pi^2 + a^2M_3)}{a^2[n^2\pi^2 + a^2M_3(1+M_1)]}, \tag{19}$$

where we observe

$$R_n^{(\mathrm{con})} < R_{n+1}^{(\mathrm{con})} . \tag{20}$$

Again the condition $W \geq 0$ in (15) gives after using $R = R_n^{(\mathrm{con})}$ as given by (19)

$$(n^2\pi^2 + a^2)^3 + N\frac{(P-1)}{(P+1)}\, n^2\pi^2 \geq 0. \tag{21}$$

For $P > 1$, the condition (21) is satisfied for all n and after noting (20) we get $R_c^{(\mathrm{con})}$, the critical value of the Rayleigh number for the convective marginal state as

$$R_c^{(\mathrm{con})} = R_1^{(\mathrm{con})} = \frac{[(\pi^2 + a^2)^3 + N\pi^2]\,(\pi^2 + a^2M_3)}{a^2[\pi^2 + a^2M_3(1 + M_1)]}. \tag{22}$$

The above result had been obtained by Das Gupta and Gupta [4]. We now consider the case $P < 1$ and determine the correct value of n from (21). We write $x_n = n^2\pi^2 + a^2$ and express (21) as

$$(x_n - \alpha_1)\ (x_n - \alpha_2)\ (x_n - \alpha_3) \geq 0, \tag{23}$$

where

$$\alpha_1 = -\left\{\frac{N(1-P)}{2(1+P)}\right\}^{1/3} [(a^2 + \beta)^{1/3} + (a^2 - \beta)^{1/3}], \tag{24}$$

$$\alpha_2 = -\left\{\frac{N(1-P)}{2(1+P)}\right\}^{1/3} [\omega(a^2 + \beta)^{1/3} + \omega^2(a^2 - \beta)^{1/3}], \tag{25}$$

$$\alpha_3 = -\left\{\frac{N(1-P)}{2(1+P)}\right\}^{1/3} [\omega^2(a^2 + \beta)^{1/3} + \omega(a^2 - \beta)^{1/3}], \tag{26}$$

and

$$\beta = \left\{a^4 - \frac{4N(1-P)}{27(1+P)}\right\}^{1/2}, \quad \omega = -\frac{1}{2}\,(1 + \sqrt{3}i). \tag{27}$$

From the above, we find that if

$$a \geq \left\{\frac{4N(1-P)}{27(1+P)}\right\}^{1/4} = a^* \quad (\text{say}), \tag{28}$$

then β is real, $0 \leq \beta \leq a^2$, α_3 is the complex conjugate of α_2 and α_1 is real and negative. We find that for all n, x_n satisfies (23). Hence, due to (20), we have $R_c^{(\mathrm{con})} = R_1^{(\mathrm{con})}$ as given in (22).

In case $a < a^*$, we put $\beta = i\beta_1$, where β_1 is obtained from (27), as

$$\beta_1 = \left\{\frac{4N(1-P)}{27(1+P)} - a^4\right\}^{1/2}. \tag{29}$$

Substituting $\beta = i\beta_1$ and expressing

$$a^2 + \beta = a^2 + i\beta_1 = \left\{\frac{4N(1-P)}{27(1+P)}\right\}^{1/2} \exp(i\psi), \tag{30}$$

$$a^2 - \beta = \left\{\frac{4N(1-P)}{27(1+P)}\right\}^{1/2} \exp(-i\psi), \tag{31}$$

where $0 < \psi < \pi/2$, we get from (24)–(26)

$$\alpha_1 = -\left[\frac{4N(1-P)}{3(1+P)}\right]^{1/2} \cos \psi/3, \tag{32}$$

$$\alpha_2 = \left[\frac{4N(1-P)}{3(1+P)}\right]^{1/2} \cos(\psi + \pi)/3, \tag{33}$$

$$\alpha_3 = \left[\frac{4N(1-P)}{3(1+P)}\right]^{1/2} \cos(\psi - \pi)/3, \tag{34}$$

and note that $\alpha_1 < 0$ and $0 < \alpha_2 < \alpha_3$. Therefore, from (23) we get the condition that

$$x_n \leq \alpha_2 \quad \text{or} \quad x_n \geq \alpha_3. \tag{35}$$

If $x_1 = \pi^2 + a^2 \leq \alpha_2$ obviously $R_c^{(\mathrm{con})}$ is given by $R_1^{(\mathrm{con})}$ which is the same as (22). If $x_1 = \pi^2 + a^2 > \alpha_2$ we have to consider all those n for which $x_n \geq \alpha_3$. Hence due to (20), we find that

$$R_c^{(\mathrm{con})} = R_{m_0}^{(\mathrm{con})}, \tag{36}$$

where

$$m_0 = \mathrm{Min}\; m\; \{n\colon x_n = n^2\pi^2 + a^2 \geq \alpha_3\}. \tag{37}$$

Obviously, $m_0 = 1$ if $\pi^2 > \alpha_3$ and $R_c^{(\mathrm{con})}$ is given by the same expression as (22). Hence we find that when $P < 1$, $a < a^*$ and $\alpha_2 - \pi^2 < a^2 < \alpha_3 - \pi^2$ (if $\pi^2 < \alpha_3$) the convective marginal state is exhibited by a perturbation with mode higher than the lowest one. In all other cases, the $R_c^{(\mathrm{con})}$ is given by $R_1^{(\mathrm{con})}$.

2.2 *Evaluation of $R_c^{(\mathrm{o.s})}$*

Using the conditions (17) we get

$$R_n^{(\mathrm{o.s})} = \frac{2(P+1)}{a^2} \frac{\left[(n^2\pi^2 + a^2)^3 + \dfrac{NP^2n^2\pi^2}{(P+1)^2}\right](n^2\pi^2 + a^2M_3)}{[n^2\pi^2 + a^2M_3(1+M_1)]} \tag{38}$$

with

$$R_n^{(\mathrm{o.s})} < R_{n+1}^{(\mathrm{o.s})}, \tag{39}$$

and after using (38), we have (as $WU + V^2 \geq 0$)

$$(n^2\pi^2 + a^2)^3 + \frac{NP^2(1-P)n^2\pi^2}{(1+3P)\,(1+P)^2} \geq 0. \tag{40}$$

Moreover from (18), we have

$$\{\sigma_n^{(i)}\}^2 = \frac{Nn^2\pi^2(1-P)}{(n^2\pi^2+a^2)\,(1+P)} - (n^2\pi^2+a^2)^2 > 0. \tag{41}$$

For $P < 1$, we find from (41) that the overstability cannot occur if

$$N < \frac{(1+P)\pi^4}{(1-P)} = N^* \quad \text{(say)}$$

and if $N > N^*$, it can occur only for those values of a for which

$$a^2 < \left[\frac{N\pi^2(1-P)}{(1+P)}\right]^{1/3} - \pi^2$$

and $R_c^{(\text{o.s})}$ is given by

$$R_c^{(\text{o.s})} = R_1^{(\text{o.s})} =$$

$$\frac{2(P+1)}{a^2}\,\frac{\left[(\pi^2+a^2)^3 + \dfrac{NP^2\pi^2}{(P+1)^2}\right](\pi^2+a^2M_3)}{[\pi^2+a^2M_3(1+M_1)]} \tag{42}$$

3. Numerical results and discussion

From the above, we observe that if $P < 1$, the convective marginal state exists and the overstability can occur only if N is sufficiently large. We have observed that in case the overstability occurs the perturbation in the marginal state must necessarily be in the lowest mode. In the case of convective marginal state, we have seen that though it can occur for all values of a, there exists a situation when the perturbation in the marginal state need not be in the lowest mode. The nature of the critical marginal state can only be determined by calculating $R_c^{(\text{con})}$ and $R_c^{(\text{o.s})}$ and finding out the smaller of the two.

In table 1, we have presented $R_c^{(\text{con})}$ for $4 \leq a^2 \leq 11.0$, $P = 0.01$, $M_1 = 10$, $M_3 = 5$, $N = 10$ and 100. For these values of the parameters overstability cannot occur and the perturbation in the marginal state is found to correspond to the lowest mode. In table 2, we have presented $R_c^{(\text{o.s})}$ and $R_c^{(\text{con})}$ for $N = 1000$ and

Table 1. Values of $R_c^{(con)}$ for $M_1=10$, $M_3=5$, for various values of a ($4.0 \le a^2 \le 11.0$) when $N=10$ and 100, $P=0.01$.

	$M_1=10,\ M_3=5$	
	$N=10$	$N=100$
a^2	$R_c^{(con)}$	$R_c^{(con)}$
4.0	89.87825	118.7332
4.5	85.88653	110.5127
5.0	82.90372	104.6493
5.5	81.11916	100.4401
6.0	80.07015	97.43691
6.5	79.58203	95.34284
7.0	79.53391	93.95316
7.5	79.83957	93.12292
8.0	80.43295	92.74441
8.5	81.27645	92.74075
9.0	82.32481	93.05116
9.5	83.55358	93.62953
10.0	84.94081	94.43945
10.5	86.46923	95.45143
11.0	88.12475	96.64412

Table 2. Values of $R_c^{(con)}$ and $R_c^{(o.s)}$ for $M_1 = 10$, $M_3 = 5$, for various values of a ($4.0 \le a^2 \le 11.0$) when $N = 1000$ and $N = 10{,}000$, $P = 0.01$.

	$M_1=10,\ M_3=5$			
	$N=1000$		$N=10{,}000$	
a^2	$R_c^{(o.s)}$	$R_c^{(con)}$	$R_c^{(o.s)}$	$R_c^{(con)}$
4.0	175.1411	6972.308	175.7124	685437.38
4.5	167.5692	5976.910	168.0650	579830.75
5.0	162.6326	5227.156	163.0634	498824.06
5.5	159.5664	4646.128	159.9490	435142.56
6.0	157.8820	4185.261	158.2258	384052.94
6.5	157.2527	3812.573	157.5647	342357.81
7.0	157.4338	3506.237	157.7392	066013.00
7.5	158.3237	3250.919	158.5867	059873.31
8.0	159.7453	3035.566	169.9889	054691.41
8.5	161.6305	2852.029	161.8574	050271.96
9.0	163.9121	2694.183	164.1244	046467.67
9.5	166.5387	2557.342	166.7382	043165.75
10.0	169.4693	2437.897	169.6574	040278.50
10.5	172.6713	2332.961	172.8492	037737.10
11.0	176.1184	2240.274	176.2871	035486.25

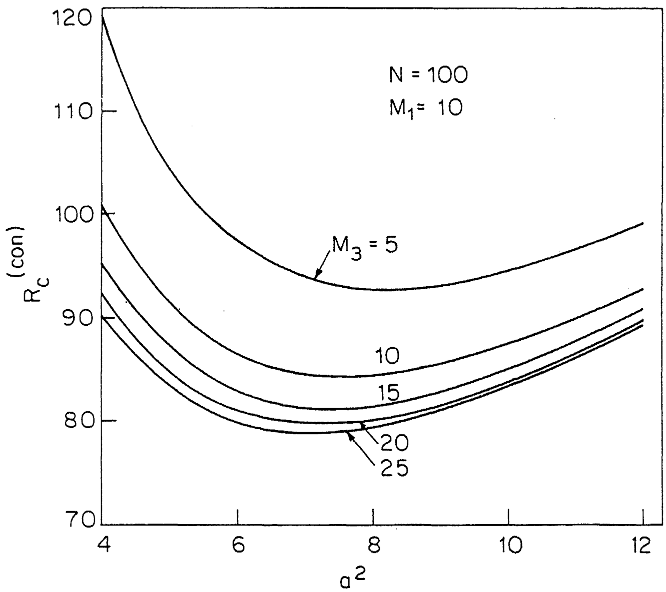


Figure 1. Plots of $R_c^{(con)}$ vs a^2 ($4.0 \le a^2 \le 12.0$) for $M_1=10$, $N=100$ and $M_3=5$, 10, 15, 20 and 25, $P=0.01$.

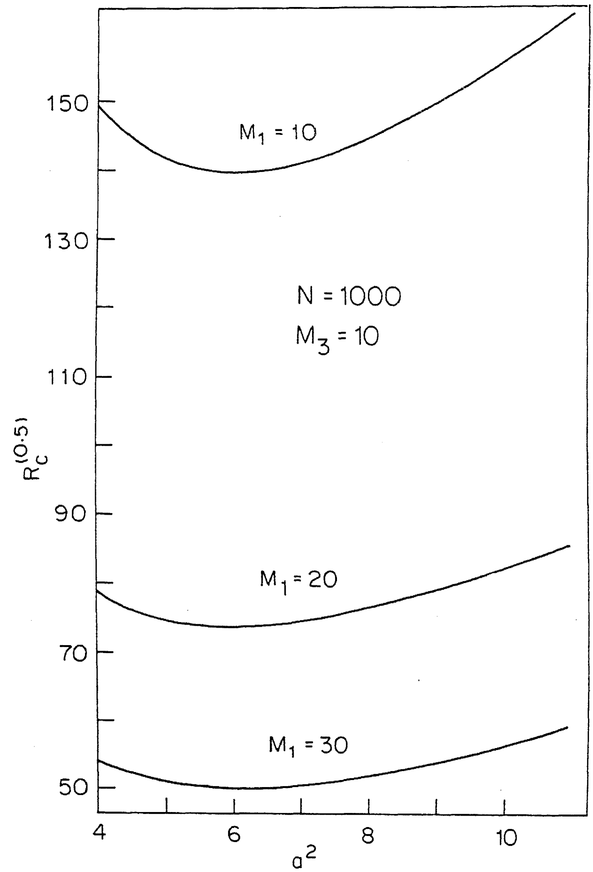


Figure 2. Plots of $R_c^{(o.s)}$ vs a^2 $(4.0 \leq a^2 \leq 11.0)$ for $M_3 = 10$, $N = 1000$ and $M_1 = 10, 20$ and 30, $P = 0.01$.

10000, other parameters being the same. We find $R_c^{(o.s)} < R_c^{(con)}$ so that the nature of the marginal state is overstable. For $R_c^{(con)}$, the perturbation has been found to correspond to higher mode. In all the above cases, the critical Rayleigh numbers increase as N increases and hence N is found to have a stabilizing effect.

We have presented the plots of $R_c^{(con)}$ vs a^2 for $P = 0.01$, $M_1 = 10$, $M_3 = 5$, 10, 15, 20 and 25 and $N = 100$ in figure 1. The plots show that M_3 has a destabilizing effect. The plots of $R_c^{(o.s)}$ vs a^2 for $P = 0.01$, $M_3 = 10$, $N = 1000$, $M_1 = 10$, 20 and 30 have been presented in figure 2. M_1 is also found to have a destabilizing effect.

References

[1] Das Gupta M and Gupta A S, Convective instability of a layer of a ferromagnetic fluid rotating about a vertical axis, *Int. J. Eng. Sci.* **17** (1979) 271–277

[2] Finlayson B A, Convective instability of ferromagnetic fluids, *J. Fluid Mech.* **40** (1970) 753–767

[3] Neuringer J L and Rosensweig R E, Ferrohydrodynamics, *Phys. Fluids* **7** (1964) 1927–1937

Proc. Indian Acad. Sci. (Math. Sci.), Vol. 95, No. 1, September 1986, pp. 31–36.

Effect of aspect ratio on the meridional circulation of a rotating homogeneous fluid

V SOMARAJU, D A MOHANDAS* and R BALASUBRAMANIAN*

S R K R Engineering College, Bhimavaram 534 202, India

*P S G College of Technology, Coimbatore 641 004, India

MS received 21 June 1985

Abstract. The effect of aspect ratio on the meridional circulation of a homogeneous fluid is analyzed. Aspect ratio is allowed to range between zero and unity. Relationships between possible horizontal and vertical length scales are obtained by length scale analysis as well as by solving an idealized problem. It is found that when $E^{1/2} \ll Z \ll E^{1/2}/\delta$, where E is the Ekman number, the stream lines are closely packed near the sidewall within a thickness of $O(E^{1/2})$. The effect of stratification is unimportant within this depth range. In the depth range $E^{1/2}/\delta \ll Z \ll 1/E\delta$ the vertical boundary layer in which the streamlines are packed is of $O(EZ\delta)^{1/3}$. When $Z \gg 1/E\delta$ it is shown that the circulation decays algebraically with depth as $1/Z$.

Keywords. Aspect ratio; meridional circulation; homogeneous fluid.

List of symbols:

u, v, w	Velocity components
ψ	Stream function
p	Pressure
H, L	Characteristic dimensions in vertical and horizontal directions
δ	aspect ratio
A_v, A_h	Vertical and horizontal eddy coefficients of momentum
E_v, E_h	Vertical and horizontal Ekman numbers
h, l	Vertical and horizontal length scales
ξ	Vertical distance
m	Fourier transform variable
f	Coriolis parameter.

1. Introduction

Analysis of sidewall friction boundary layers has assumed considerable importance since these layers play a significant part in ocean circulation, spin-up problems etc.

Stewartson [6] considered steady linear axisymmetric flow of a homogeneous fluid in a rotating cylindrical container with unit aspect ratio and found that the meridional circulation of fluid driven by the Ekman layers is completed within two sidewall layers, now known as Stewartson layers of thickness $E^{1/4}$ and $E^{1/3}$ where $E(=\nu/\Omega L^2)$ is the Ekman number. Pedlosky [4] and Durance and Johnson [3] analyzed the upwelling boundary layers whose axial scale is $O(1)$ for a linear and homogeneous oceanic model with β-plane approximation. While Pedlosky's analysis spanned the range for the aspect ratio $\delta \ll E^{1/2}$, the analysis of Durance and Johnson was for the range $\delta \gg E^{1/2}$. The aim of this paper is to understand the role of various horizontal and vertical length scales associated with meridional circulation in a rotating homogeneous fluid with a constant Coriolis parameter for $0 \ll \delta < O(1)$.

The main motivation of this work is to elucidate the effect of aspect ratio on the various length scales and circulation pattern in a rotating hydrodynamic flow and to draw comparisons with other studies wherever possible. To keep the model more akin to an oceanic model, our mathematical analysis closely follows that of Blumsack [2] who had analyzed the transverse circulation near a coast in a rotating stratified fluid for $\delta^2 \ll S \ll 1$ where S is the stratification parameter. Since Blumsack's work was concerned with a stratified fluid and $S \gg \delta^2$ the effect of aspect ratio on the length scales and circulation pattern has not received considerable attention.

2. Governing equations

Following Blumsack [2], we assume that the perturbation state is independent of y, the alongshore coordinate and all dependent variables are functions of the offshore coordinate x and the vertical coordinate z. Then the linear steady state equations in nondimensional form (for example see [2,4]) are,

$$
\begin{aligned}
-v &= -\frac{\partial P}{\partial x} + \frac{E_V}{2}\frac{\partial^2 u}{\partial z^2} + \frac{E_H}{2}\frac{\partial^2 u}{\partial x^2}, \\
u &= \frac{E_V}{2}\frac{\partial^2 v}{\partial z^2} + \frac{E_H}{2}\frac{\partial^2 v}{\partial x^2}, \qquad\qquad (1) \\
o &= -\frac{\partial P}{\partial z} + \delta^2\left[\frac{E_V}{2}\frac{\partial^2 w}{\partial z^2} + \frac{E_H}{2}\frac{\partial^2 w}{\partial x^2}\right], \\
o &= \frac{\partial u}{\partial x} + \frac{\partial w}{\partial z},
\end{aligned}
$$

Here δ = vertical length scale divided by horizontal length scale = H/L is an aspect ratio; u,v,w are velocity components, P is the pressure, $E_V = 2A_V/fH^2$ and $E_H = 2A_H/fL^2$ are the vertical and horizontal Ekman numbers where A_V and A_H are eddy coefficients of momentum and f is the Coriolis parameter.

Introducing stream function ψ so that $u=-\partial\psi/\partial z$ and $W=\partial\psi/\partial x$ and eliminating P and V from equations, we get

$$\left(\delta^2\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial z^2}\right)\left(\frac{E_V}{2}\frac{\partial^2}{\partial z^2}+\frac{E_H}{2}\frac{\partial^2}{\partial x^2}\right)^2\psi+\frac{\partial^2\psi}{\partial z^2}=0.$$

In oceanic models, since it is difficult to determine E_H and E_V, it is assumed that $E_V=E_H=E$ for simplicity of exposition. This assumption leads to $H=L(A_V/A_H)^{1/2}$. The above equation becomes,

$$\frac{\partial^2\psi}{\partial z^2}+\frac{E^2}{4}\nabla^4\left(\frac{\partial^2}{\partial z^2}+\delta^2\frac{\partial^2}{\partial x^2}\right)\psi=0, \tag{2}$$

where

$$\nabla^2=\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial z^2}, \quad E=2A_H/fL^2.$$

3. Length scale analysis

To find the length scale relationships, replace in the above equation $\partial/\partial x$ by l^{-1} and $\partial/\partial z$ by h^{-1} where l and h are the horizontal and vertical length scales respectively. Equation (2) then becomes,

$$h^4l^6+\frac{E^2}{4}(h^2+l^2)^2(\delta^2h^2+l^2)=0. \tag{3}$$

When $l\gg h$, equation (3) gives

$$h\approx E^{1/2} \text{ if } l\gg E^{1/2}.$$

This vertical length scale corresponds to the Ekman depth.

For $h\gg l$, (3) becomes

$$h^4l^6+\frac{1}{4}E^2h^4l^2+\frac{1}{4}E^2\delta^2h^6=0. \tag{4}$$

From (4) we get

$$l\approx\delta h \text{ if } h\ll E^{1/2}/\delta, \tag{5a}$$

$$l\approx E^{1/2} \text{ if } E^{1/2}\ll h\ll E^{1/2}/\delta, \tag{5b}$$

$$l\approx(Eh\delta)^{1/3} \text{ if } h\gg E^{1/2}/\delta. \tag{5c}$$

For $h=O(1)$, these length scales agree with those discussed earlier (see [4] and [3]). It may be noted that the hydrostatic $E^{1/3}$ layer (see [4]) which arises solely due to β-effect does not appear in our case due to the assumption of constant coriolis parameter. For $h\gg E^{1/2}/\delta$, the nonhydrostatic horizontal length scale $(Eh\delta)^{1/3}$ increases with h. If $\delta=O(1)$, we have $l\approx(Eh)^{1/3}$ for $h\gg E^{1/2}$ and therefore its

thickness will be $O(E^{1/3})$ when $h=O(1)$. If $h=O(1)$ and $\delta \gg E^{1/2}$ the thickness of this layer becomes $(E\delta)^{1/3}$. It may be guessed that this layer will be dynamically similar to Stewartson's $E^{1/3}$ layer, and it has been analyzed by Durance and Johnson [3].

From (5a,b,c), for $h \ll E^{1/2}/\delta$ we see that the $(Eh\ \delta)^{1/3}$ layer splits into $E^{1/2}$ and $h\delta$ layers. These length scales arise due to the constraint that $\delta \ll 1$. If $h=O(1)$, these vertical layers exist for $\delta \ll E^{1/2}$. If $\delta = O(1)$, we see from (5) that these length scales merge together and correspond to the corner region whose horizontal and vertical length scales are $O(E^{1/2})$.

In what follows, we shall consider an infinite depth fluid and see how the various length scales mentioned above manifest themselves in the circulation pattern. To achieve this, the integral Fourier transform technique will be used, and to suit this technique, the boundary conditions will be assumed suitably.

4. Structure of the boundary layers

We consider an infinite depth fluid and assume that there is a stress $F(x)$ acting only in the y-direction (along shore) on the fluid horizontal surface at $z=0$. $F(x)$ is supposed to tend to zero as $x \to \infty$. The boundary conditions at $z=0$ may then be written as

$$V_z(x,0) = F(x);$$

$$\psi(x,0) = \psi_{zz}(x,0) = 0.$$

The boundary conditions at $x=0$ are all taken as homogeneous. These are,

$$V_z = \psi = \psi_{xx} = 0 \quad \text{at } x=0.$$

These conditions are appropriate for a problem in which $F(x)$ is an odd function and we shall be describing a shear layer at $x=0$ if $F(O) \neq 0$. However, as noted by Blumsack [2], these boundary conditions can also be extended to the case of a rigid wall (where the correct condition is $\psi_x\ (0,z)=0$ instead of $\psi_{xx}(0,z)=0$) without affecting the physics of the problem if the boundary condition on w is satisfied in a region too thin to affect the circulation pattern.

Fourier sine transform analysis of (2) subject to the above boundary conditions gives for values of $\xi(=-z)$ much larger than the Ekman depth $E^{1/2}$

$$\psi\ (x,\xi) = \frac{E}{\pi} \int_0^\infty \tilde{F}(m) \frac{\sin\ mx}{1+\frac{E^2}{4}\ m^4}\ \exp\ (-\xi/Z_D)\,\mathrm{d}m, \tag{6}$$

where

$$Z_D(m) = \left(1+\frac{E^2}{4}\ m^4\right)^{1/2} \Big/ \frac{E}{2}\ m^3\,\delta$$

and $\tilde{F}(m)$ is the Fourier sine transform of $F(x)$ defined as

$$\tilde{F}(m) = \int_0^{\infty} F(x) \sin mx \, dx$$

For $E^{1/2} \ll \xi \ll E^{1/2}/\delta$ we see that $Z_D(m) \gg \xi$ for $m \leq O(E^{-1/2})$.

Hence in this depth range, $\exp(-\xi/Z_D)$ in (6) may be replaced by unity. Assuming $F(x) = e^{-x}$, we get from (6)

$$\psi(x,\xi) = \frac{E}{2} e^{-x} \left[1 - \exp(-xE^{1/2}) \cos(xE^{-1/2}) \right]. \tag{7}$$

Thus there is a boundary layer of thickness $O(E^{1/2})$ at $x = 0$ within which there is an intense vertical flow. This layer serves to complete the meridional circulation induced due to Ekman pumping. (The nonhydrostatic thinner layer of thickness $O(\delta h)$ carries negligible mass flux.) We see from (7) that for $F(x)$ positive and $F'(x)$ (which gives the curl of the wind stress) negative, the Ekman layer pumps the fluid into the interior and the circulation is closed by an intense upward motion in the $E^{1/2}$ vertical layer. This result is physically obvious if we recall (see [5]) that the horizontal mass flux associated with the free surface Ekman layer friction velocity is completely perpendicular to the applied stress and the interior vertical velocity at the lower edge of the Ekman layer is given by the curl of the wind stress. Further this solution is the same as that of Blumsack [2] obtained for the stratified case. The effect of stratification is not felt in the above depth range.

Consider now depths in the range $E^{1/2}/\delta \ll \xi \ll 1/E\delta$. For $\xi \gg E^{1/2}/\delta$ it is easily seen from the expression for $Z_D(m)$ that only values of $m \ll E^{-1/2}$ contribute significantly to ψ in (6). Equation (6) becomes

$$\psi(x,\xi) = \frac{E}{\pi} \int_0^{\infty} \frac{m}{1+m^2} \sin(mx) \exp\left[-(m^3 \delta E)\xi \right] dm. \tag{8}$$

As such, it seems difficult to obtain a closed form solution for the above integral when $E\delta\xi \leq O(1)$. However, we see from the equation, that near $x = 0$, there is a region of width $(E\delta\xi)^{1/3}$ in which streamlines are packed. The width of the region varies from $E^{1/2}$ when $\xi = E^{1/2}/\delta$ to $O(1)$ when $\xi = 1/E\delta$. Thus the upwelling region whose radial extent remains constant upto a depth of $E^{1/2}/\delta$ spreads laterally for greater depths.

Finally, we investigate the circulation for $\xi \gg (E\delta)^{-1}$. Evaluating (8) by Laplace's method [1], we get,

$$\psi(x,\xi) \approx \frac{x}{3\pi\delta\xi}.$$

Hence, the transverse circulation decays with depth algebraically as $1/\xi$. It may be noted that in the stratified case [6], it decays like $1/\xi^3$. Since the stratification inhibits the vertical flow, the decay is faster in the stratified case than in the homogeneous case. Also we see that, for homogeneous case, the upwelling layer

extends down to a greater depth and it increases as the aspect ratio decreases. Further, since the vertical mass flux decreases with depth, none may enter the lower Ekman Layer. This justifies the assumption of infinite depth fluid.

Acknowledgement

This work has been partially supported by the University Grants Commission, New Delhi.

References

[1] Bender C M and Orszag S A (1978) Advanced Mathematical methods for Scientists and Engineers, McGraw-Hill Book Company
[2] Blumsack S (1972) The transverse circulation near at coast, *J. Phys. Oceanography*, **2** 34–40
[3] Durance J A and Johnson J A (1970) East Coast Ocean currents, *J. Fluid Mech.* **44** 161–172
[4] Pedlosky J (1968) An overlooked aspect of the wind-driven oceanic circulation, *J. Fluid Mech.* **32** 809–821
[5] Pedlosky J (1979) *Geophys. Fluid Dynamics* Springer Verlag
[6] Stewartson K (1957) On almost rigid rotations, Part 2, *J. Fluid Mech.* **3** 17–26

Proc. Indian Acad. Sci. (Math. Sci.), Vol. 95, No. 1, September 1986, pp. 37–40.

Analytic and harmonic maps into a topological space

S ARUNDHATHI and S H KULKARNI

Department of Mathematics, Indian Institute of Technology, Madras 600 036, India

MS received 8 May 1985; revised 14 April 1986

Abstract. The relationship between the harmonicity and analyticity of a continuous map from the open unit disc to the underlying space of a real algebra is investigated.

Keywords. Analytic map; harmonic map; function algebra.

1. Introduction

Real function algebras were defined in [3] and certain aspects of the theory of real function algebras were developed along the lines of complex function algebras. In particular, sufficient conditions were given for the presence of an analytic structure in the carrier space of a real function algebra, [3]. These conditions automatically imply harmonic structure both in the carrier space as well as the maximal ideal space. A question was raised whether weaker conditions could be found which would imply harmonic structure in the maximal ideal space without necessarily implying analytic structure in the carrier space. This question acquires greater significance in view of the result that in a complex function algebra the existence of a harmonic structure implies that of an analytic structure or an anti-analytic structure, [4]. So, it is natural to inquire whether a similar relationship holds in a real function algebra.

In this paper, we prove the following two results which provide partial answers to the above question. In fact the first result has been proved in a more general setting than that of a real function algebra.

1. Let X be a Hausdorff topological space and A, a real algebra of continuous complex-valued functions defined on X. Let U be the open unit disc in the complex plane, $F: U \to X$ a continuous map such that Re $(f \circ F)$ is harmonic for all f in A and Y be the set $Y = \{x \in X : \operatorname{Im} g(x) = 0 \text{ for all } g \text{ in } A\}$. Then $f \circ F$ is analytic for all f in A or $\overline{f \circ F}$ is analytic for all f in A on each connected component of $U - F^{-1}(Y)$.

2. If X is a compact plane set whose interior is connected and which is symmetric with respect to the real axis, τ the complex conjugation and if A is a real function algebra on (X, τ) such that Ref is harmonic in $X^0 =$ interior of X for all f in A then f is analytic for all f in A or f is anti-analytic for all f in A in X^0.

2. Basic definitions and lemmas

As usual $\mathbf{R}$ denotes the real line, $\mathbf{C}$ denotes the complex-plane and U the open unit disc in $\mathbf{C}$.

Let K be a non-empty open subset of $\mathbf{C}$, f a complex-valued function on K such that the real and imaginary parts of f have continuous partial derivatives of order k. Such functions will be referred to as C^k-maps. We recall that f is said to be anti-analytic if $\bar{f}$ is analytic.

LEMMA 1. Let f be a complex-valued function defined on an open connected subset K of $\mathbf{C}$ whose real and imaginary parts u and v are C^2-maps. Assume that $\mathrm{Re}f$, $\mathrm{Re}f^2$, $\mathrm{Re}f^3$ and $\mathrm{Re}f^4$ are all harmonic functions. Then, f is an analytic function on K or an anti-analytic function on K.

Proof. Refer [4], theorem 2.2.

LEMMA 2. Let K be an open connected set in $\mathbf{C}$ and A a subset of $C(K)$, the set of all continuous, complex-valued functions on K. Suppose A is closed under multiplication, and every function in A is analytic or anti-analytic in K. Then all functions in A are analytic or all functions in A are anti-analytic.

Proof. Proof is elementary and hence omitted. Readers can refer to theorem 3-1 of [4].

LEMMA 3. Let V be an open set in $\mathbf{R}^n$, A a ring of continuous complex-valued functions (under point-wise multiplication) on V such that $\mathrm{Re}f$ is of class C^k for each $f \in A$. Let $Y = \{x \in V : \mathrm{Im}g(x) = 0 \text{ for all } g \text{ in } A\}$. Then $\mathrm{Im}f$ is of class C^k on $V - Y$ for each f in A.

Proof. Let $f = u + iv \in A$ and $x \in V - Y$. Then there exists g in A with say $g = u_1 + iv_1$ such that $v_1(x) \neq 0$.

Since v_1 is continuous in V, we can find a neighbourhood D of x such that $v_1(x) \neq 0$ for all x in D.

By hypothesis, $u = \mathrm{Re}f$, $u_1 = \mathrm{Re}g$ and $\mathrm{Re}(fg) = uu_1 - vv_1$ are of class C^k. Thus vv_1 is of class C^k.

But in D, $v = vv_1/v_1$. Both vv_1 and v_1 are of class C^k in D and v_1 is non-vanishing in D. Hence v is of class C^k in D and in particular at x. Thus, v is a C^k-map in $V - Y$.

3. Main theorems

Let X be a Hausdorff topological space and A a real algebra of continuous, complex-valued functions on X. A continuous map $F: U \to X$ is called a harmonic map if $\mathrm{Re}(f \circ F)$ is harmonic on U for all f in A. F is called an analytic map

(respectively an anti-analytic map) if the function $f \circ F$ is analytic (respectively anti-analytic) on U for all f in A.

THEOREM 4. Let X be a Hausdorff topological space and A a real algebra of continuous complex-valued functions defined on X. Let $Y = \{x \in X : \mathrm{Im} g(x) = 0$ for all g in $A\}$ and $F: U \to X$ a harmonic map. Then F is an analytic map or an anti-analytic map on each connected component of $U - F^{-1}(Y)$.

Proof. Let $\tilde{f} = f \circ F$ for f in A, and $\tilde{A} = \{\tilde{f} : f \in A\}$. Then, $\tilde{A}$ is a real algebra on U. Since $\mathrm{Re}\tilde{f} = \mathrm{Re}(f \circ F)$ is harmonic, it is of class C^2 on U for each $\tilde{f} \in \tilde{A}$.

Hence by lemma (3), $\mathrm{Im}\tilde{f}$ is of class C^2 on $U - F^{-1}(Y)$ for all $\tilde{f}$ in $\tilde{A}$. Now let K be a connected component of $U - F^{-1}(Y)$. Then it is easy to see that each $\tilde{f} \in \tilde{A}$ satisfies the hypotheses of lemma (1).

Hence $\tilde{f}$ is analytic or anti-analytic by lemma (2). Therefore, by lemma (3) all functions in $\tilde{A}$ are analytic on K or all functions in $\tilde{A}$ are anti-analytic on K.

COROLLARY 5. If $U - F^{-1}(Y)$ is connected then F is analytic or anti-analytic in $U - F^{-1}(Y)$. In particular, if $F^{-1}(Y)$ is the null set then F is an analytic map or an anti-analytic map in U.

4. Real function algebras

Let X be a compact, Hausdorff space. By $C(X)$ (respectively $C_{\mathbf{R}}(X)$) we denote the complex (respectively real) Banach algebra of all continuous, complex-valued (respectively real-valued) functions on X with the supremum norm. A homeomorphism $\tau: X \to X$ with $\tau^2 = \tau \circ \tau =$ identity mapping on X is called an involution on X or an involutionary homeomorphism on X. Let

$$C(X, \tau) = \{f \in C(X) \colon f(\tau(x)) = \overline{f(x)} \text{ for all } x \text{ in } X\}.$$

Then $C(X, \tau)$ is a real commutative Banach algebra with the identity 1. Also $C(X, \tau)$ separates points on X, that is, for any x_1, x_2 in X with $x_1 \neq x_2$, there exists $f \in C(X, \tau)$ such that $f(x_1) \neq f(x_2)$. A real function algebra on (X, τ) is a real subalgebra A of $C(X, \tau)$ such that (i) A is uniformly closed in $C(X, \tau)$; (ii) A contains the real constants; (iii) A separates points on X. For examples and other details on complex function algebras refer [1] and [2] and for details about real function algebras, see [3].

Remark 6. Let A be a real function algebra on (X, τ) and $Y = \{x \in X : \tau(x) = x\}$ the set of fixed points of τ. Since A separates the points of X, it is easy to see that $Y = \{x \in X\colon \mathrm{Im} g(x) = 0$ for all g in $A\}$. Thus, in view of theorem (4) if $F: U \to X$ is a continuous map such that $\mathrm{Re}(f \circ F)$ is harmonic for all f in A then F is an analytic or anti-analytic map in the connected components of $U - F^{-1}(Y)$.

Example 7. Let X be the annular region in $\mathbf{C}$ defined by $X=\{z\in\mathbf{C}:r\leq|z|\leq 1/r\}$ for some r with $0<r<1$ and let the involutionary homeomorphism τ be given by the equation $\tau(z)=-1|\bar{z}$ for all z in X. Let A be a real function algebra on (X,τ). Note that τ has no fixed points in X. By corollary 5 if $F: U\to X$ is a harmonic map then F must be an analytic map or an anti-analytic map on U.

Note 8. In the case of a complex function algebra every harmonic map is an analytic map or an anti-analytic map, (theorem 3.1 of [4]), whereas in the case of a real (or a real function) algebra we could only prove that a harmonic map is an analytic map or an anti-analytic map only in connected components of $U-F^{-1}(Y)$. However, we point out that under the hypotheses of theorem 4 the set $F^{-1}(Y)$ has an empty interior if F is non-constant. (Note that if F is constant, it is trivially an analytic map). For, if $F^{-1}(Y)$ contains a disc D then for f in A and $\tilde{f}=f\circ F=u+iv$, $v=0$ throughout D.

But v^2 is real analytic, hence is identically 0 in U. Thus u and u^2 are harmonic in U, hence $\tilde{f}=u$ is constant, a contradiction.

THEOREM 9. Let X be a subset of $\mathbf{C}$ which is compact with connected interior and which is symmetric with respect to the real-axis, that is $\bar{z}\in X$ whenever $z\in X$ and $\tau: X\to X$ be defined by $\tau(z)=\bar{z}$ for all z in X. Suppose that A is a real function algebra on (X,τ) such that Ref is harmonic in $X^0=$ interior of X for all f in A. Then f is analytic for all f in A or anti-analytic for all f in A in X^0.

Proof. Let $K=\{z\in X^0:\operatorname{Im} z>0\}$ and $f=u+iv$ be in A.

By hypotheses, Ref, Ref^2, Ref^3, and Ref^4 are all harmonic in K. Hence by lemma 3 $v=\operatorname{Im} f$ is of class C^2 and since K is connected, lemma 1 shows that f is analytic or anti-analytic in K.

Since $f(\bar{z})=\overline{f(z)}$ for all z in X, by Schwarz's reflection principle f is analytic or anti-analytic in the whole of X^0. Now lemma 2 implies that f is analytic for all f in A in X^0 or f is anti-analytic for all f in A in X^0.

Acknowledgement

The authors thank the referee for several suggestions to improve this paper.

References

[1] Browder A, *Introduction to function algebras*, W. A. Benjamin, New York (1969)
[2] Gamelin T W, *Uniform algebras*, Prentice-Hall, Englewood Cliffs, New Jersey (1969)
[3] Kulkarni S H and Limaye B V, Gleason parts of real function algebras, *Can. J. Math.* **33** (1981) 181–200
[4] Kulkarni S H, Analytic and harmonic maps in the maximal ideal space of a function algebra, *J. Math. Phys. Sci.* Indian Inst. Technol. Madras **17** (1983) 169–175

Proc. Indian Acad. Sci. (Math. Sci.), Vol. 95, No. 1, September 1986, pp. 41–44.

On subnormal operators

B C GUPTA

Department of Mathematics, Sardar Patel University, Vallabh Vidyanagar 388 120, India

MS received 10 September 1985

Abstract. Let S be a pure subnormal operator such that $C^*(S)$, the C^*-algebra generated by S, is generated by a unilateral shift U of multiplicity 1. We obtain conditions under which S is unitarily equivalent to $\alpha+\beta U$, α and β being scalars or S has C^*-spectral inclusion property. It is also proved that if in addition, S has C^*-spectral inclusion property, then so does its dual T and $C^*(T)$ is generated by a unilateral shift of multiplicity 1. Finally, a characterization of quasinormal operators among pure subnormal operators is obtained.

Keywords. Subnormal operators; self-dual subnormal operators; quasinormal operators; unilateral shifts; C^*-algebra; C^*-spectral inclusion property.

By an operator we mean a bounded linear operator on a fixed separable infinite dimensional Hilbert space H. For an operator A, let $\sigma_e(A)$, $\sigma_a(A)$ and $\sigma(A)$ denote respectively the essential spectrum, the approximate point spectrum and the spectrum of A and let $C^*(A)$ denote the C^*-algebra generated by A. Let S be a subnormal operator on H and

$$N=\begin{pmatrix} S & X \\ o & T^* \end{pmatrix} \ldots. \quad (*)$$

be its minimal normal extension on $\mathcal{K}=H\oplus H^{\perp}$. The subnormal operator T appearing in the representation $(*)$ is called the dual of S and has been studied by Conway [5]. Dual of S is unique up to unitary equivalence; and if S is pure, then N^* is the minimal normal extension of T. A subnormal operator S is said to be self-dual if it is unitarily equivalent to its dual.

A subnormal operator S is said to have C^*-spectral inclusion property (C^*-SIP) if $\sigma(L)\subset\sigma(T_L)$ for every operator $L\in C^*(N)$, where T_L is the compression of L to H. Keough [7] has shown that S has C^*-SIP if and only if $\sigma(N)=\sigma_a(S)$.

It is known that every pure quasinormal operator and so every unilateral shift of multiplicity 1 is self-dual and has C^*-SIP [5], [8]. In this paper, we investigate the properties of pure subnormal operators S having C^*-SIP and for which $C^*(S)$ is generated by a unilateral shift of multiplicity 1.

One may ask if there are self-dual pure subnormal operators S other than translations of scalar multiples of unilateral shift of multiplicity 1 (up to unitary equivalence) such that S has C^*-SIP and $C^*(S)$ is generated by a unilateral shift of

multiplicity 1. For pure quasinormal operators, as we shall show below, the answer is no. We need the following result which appeared in [6].

THEOREM 1. Let T be a hyponormal operator. Then $C^*(T)$ is generated by a unilateral shift of multiplicity 1 if and only if T is unitarily equivalent to an operator A satisfying the following properties:

(a) A is irreducible;
(b) A^*A-AA^* is compact;
(c) $\sigma_e(A)$ is a simple closed curve γ;
(d) $\sigma(A)=\gamma \cup G$, where G is the bounded component of the complement of γ;
(e) for $\lambda \in G$, $\operatorname{ind}(A-\lambda)=-1$.

Let θ denote the class of all operators A for which A^*A commutes with $A+A^*$. Every quasinormal operator belongs to θ, and every hyponormal operators in θ is subnormal [3]. It is not known whether these subnormal operators are self-dual or have C^*SIP. However, we prove the following.

THEOREM 2. Let S be a self-dual pure subnormal operator in θ. If $C^*(S)$ is generated by a unilateral shift U of multiplicity 1, then S is unitarily equivalent to $\alpha+\beta\, U$ for some scalars α and β.

First we establish the following particular case which appears to be of independent interest.

PROPOSITION. Let S be a pure quasinormal operator. If $C^*(S)$ is generated by a unilateral shift of multiplicity 1, then S is unitarily equivalent to a constant multiple of U.

Proof. By Brown's characterization of quasinormal operators obtained in [1], there is a positive definite operator P on a Hilbert space M such that S is unitarily equivalent to

$$\begin{pmatrix} o & & & & & \\ P & o & & & & \\ o & P & o & & & \\ & o & P & o & & \\ & & \cdot & \cdot & \cdot & \\ & & & \cdot & \cdot & \cdot \\ & & & & \cdot & \cdot & \cdot \end{pmatrix}$$

on $M \oplus M \oplus M \ldots.$ Since $C^*(S)$ is generated by a unilateral shift of multiplicity 1, S satisfies the conditions (a), (b) and (c) of theorem 1. Therefore P must be compact and by corollary 2 of [11], $\sigma(P)=\{c\}$ for some constant $c>0$. The irreducibility of S now gives the desired result.

Proof of theorem 2. Since S is self-dual, by a result of Murphy [10], there exists a normal operator B such that $[S]=S^*S-SS^*=BB^*$ and $S^*B=BS$. Since $S\in\theta$, we have $S^*BB^*=BB^*S$. Therefore $B^*SB=S^*B^*B=BB^*S=B^*S^*B$ and so $[S]S[S]$ is self-adjoint. Now by theorem 4 of [2], $S=A+Q$ where A is a self-adjoint operator commuting with every operator in $C^*(S)$ and Q is a quasinormal operator. Since $C^*(S)$ is generated by the unilateral shift U of multiplicity 1, the self-adjoint operator A must be unitarily equivalent to an analytic Toeplitz operator on the H^2-space of unit circle [4, corollary 6.13]. Therefore $A=\alpha I$ for some scalar α, $S-\alpha I$ is pure quasinormal and $C^*(S-\alpha I)$ is generated by U. The above proposition now implies that there is a scalar β such that S is unitarily equivalent to $\alpha+\beta S$.

In [7], Keough has shown that if a pure subnormal operator S has C^*-SIP, then its minimal normal extension N has no isolated eigenvalues of finite multiplicity. In the converse direction, we have the following.

THEOREM 3. Suppose S is a self-dual subnormal operator and $C^*(S)$ is generated by a unilateral shift of multiplicity 1. If the minimal normal extension N has no isolated eigenvalues of finite multiplicity, then S has C^*-SIP.

Proof. Since $C^*(S)$ is generated by unilateral shift of multiplicity 1, S satisfies conditions (a) to (d) of theorem 1. The dual T of S is unitarily equivalent to S and so $\sigma(S)=\gamma\cup G$ is symmetric about the real axis [5, Proposition 2.1]. Therefore $\sigma_e(S)=\gamma$ is also symmetric about the real axis. Using theorem 13.9 of [4], we have $\sigma_e(N)=\sigma_e(S)\cup\sigma_e(T^*)=\sigma_e(S)\cup\sigma_e(S^*)=\sigma_e(S)=\sigma_a(S)$; and since N has no isolated eigenvalues of finite multiplicity, $\sigma(N)=\sigma_e(N)$. Therefore $\sigma(N)=\sigma_a(S)$ as required.

Our next result shows that a subnormal operator S with C^*-SIP for which $C^*(S)$ is generated by a unilateral shift of multiplicity 1 strives to be a self-dual.

THEOREM 4. If a subnormal operator S has C^*-SIP and $C^*(S)$ is generated by a unilateral shift of multiplicity 1, then its dual T has C^*-SIP and $C^*(T)$ is generated by a unilateral shift of multiplicity 1. If in addition, $\sigma_e(S)$ is symmetric about the real axis, then there exists a unitary operator V and a compact operator K such that $T=VSV^{-1}+K$.

Proof. The operator S satisfies conditions (a) to (e) of theorem 1. By theorem 13.9 of [4], corollary 1.8 and proposition 1.4 of [5] we have (i) $\sigma_a(S)=\sigma_e(S)$, (ii) T is irreducible, (iii) T^*T-TT^* is compact, (iv) $\sigma_a(T)=\sigma_e(T)$ and (v) $\sigma(T)=\sigma(S^*)$.

We first show that $C^*(T)$ is generated by a unilateral shift of multiplicity 1. In view of the facts (ii), (iii) and (v) and theorem 1, it is sufficient to show that $\sigma_e(T)=\gamma^*$, the complex conjugate of γ; and ind $(T-\lambda)=-1$ for every $\lambda\in G^*$.

Since S has C^*-SIP, N has no isolated eigenvalues of finite multiplicity. Therefore $\sigma_e(N)=\sigma(N)=\sigma_a(S)=\sigma_e(S)$; and since $\sigma_e(N)=\sigma_e(S)\cup\sigma_e(T^*)$, we must have $\sigma_e(T^*)\subset\sigma_e(S)=\gamma$. Now for $\lambda\in G$, $0=\operatorname{ind}(N-\lambda)=\operatorname{ind}(S-\lambda)+\operatorname{ind}$

$(T^*-\lambda) = -1+\text{ind}(T^*-\lambda)$ so that $\text{ind}(T^*-\lambda)=1$. Therefore $\sigma_e(T^*)$ must separate the plane and so we have $\sigma_e(T^*)=\gamma$, that is, $\sigma_e(T)=\gamma^*$. Now if $\lambda \in G^*$, then $\text{ind}(T-\lambda) = -\text{ind}(T^*-\lambda^*) = -1$.

Since $\sigma(N^*)=\gamma^*=\sigma_e(T)=\sigma_a(T)$, T has C^*-SIP.

The last conclusion of the theorem follows from theorem 11.1 of [9].

If S is a quasinormal operator with $\ker S=\{0\}$ and if N is the minimal normal extension of S, then $H \subset \overline{N^*(H)}$ [8]. Since S^*S commutes with SS^*, this, when combined with the following result, characterizes quasinormal operators among pure subnormal operators.

THEOREM 5. If S is a pure subnormal operator for which S^*S commutes with SS^* and $H \subset \overline{N^*(H)}$, then S is quasinormal.

Proof. Consider the matrix representation $(*)$ of N. Let h be an arbitrary vector in H and let

$$k=\begin{pmatrix} S^*[S]h \\ X^*S^*Sh \end{pmatrix}.$$

Now $SS^*[S]h=[S]SS^*h=XX^*SS^*h$. Therefore $Nk \in H^\perp$, that is, $k \perp N^*(H)$. Since $H \subset \overline{N^*(H)}$, $k \in H^\perp$ and so $S^*[S]h=0$. Therefore $S^*[S]=0$ and S is quasinormal.

Acknowledgement

The author gratefully acknowledges financial support from University Grants Commission, New Delhi.

References

[1] Brown A, On a class of operators, *Proc. Am. Math. Soc.* **4** (1953) 723–728
[2] Campbell S L, Linear operators for which T^*T and $T+T^*$ commute, *Pac. J. Math.* **61** (1975) 53–57
[3] Campbell S L, Linear operators for which T^*T and $T+T^*$ commute-III, *Pac. J. Math.* **76** (1978) 17–19
[4] Conway J B, *Subnormal operators*, Pitman Advanced Publishing Program 1981
[5] Conway J B, The dual of a subnormal operator, *J. Operator Theory* **5** (1981) 195–211
[6] Conway J B and McGuire P, Operators with C^*-algebra generated by a unilateral shift, *Trans. Am. Math. Soc.* **284** (1984) 153–161
[7] Keough G E, Subnormal operators, Toeplitz operators, and Spectral inclusions, *Trans. Am. Math. Soc.* **263** (1981) 125–135
[8] Keough G E, *Subnormal operators, Toeplitz operators and spectral inclusion theorem*, Ph.D. dissertation, Indiana University 1979
[9] *Lecture notes in mathematics*, No. 345, Springer Verlag 1973.
[10] Murphy G J, Self-dual subnormal operators, *Comment Math. Univ. Carolin.* **23** (1982) 467–473
[11] Williams L R, Equality of essential spectra of quasinormal operators, *J. Oper. Theor.* **3** (1980) 57–69

Proc. Indian Acad. Sci. (Math. Sci.), Vol. 95, No. 1, September 1986, pp. 45–51.

Computation of $g(1,5;\ 10)$

VASANTI N BHAT-NAYAK and ANJANA WIRMANI-PRASAD

Centre of Advanced Study in Mathematics, University of Bombay, Vidyanagari, Bombay 400 098, India

MS received 10 December 1985

Abstract. It is established that the exact covering number $g(1,5;\ 10)$ is 102. It is further shown that this configuration is unique. It can be obtained from the unique Steiner system $S(5,\ 6,\ 12)$.

Keywords. Exact covering number; Steiner system.

1. Introduction

A $(\lambda, \mu; v)$ design is an arrangement of v varieties (also called points or elements) into blocks in such a way that every set of μ varieties occurs in exactly λ blocks $(\lambda \geqslant 1, 1 \leqslant \mu < v)$. The blocks must be incomplete (length $v-1$ or less), but need not all have the same length. The $\lambda - \mu$ problem is to determine $g(\lambda, \mu; v)$, the minimum number of blocks in any $(\lambda, \mu; v)$, design. A $(\lambda, \mu; v)$ design with $g(\lambda, \mu; v)$ blocks is called *minimal*. A Steiner system $S(t, k, v)$ is arrangement of v treatments into blocks of size k such that every t-tuple occurs exactly once.

In this paper we show that $g(1,\ 5;\ 10) = 102$ and that there is a unique configuration achieving this bound. It is known [1] that Steiner system $S(5,\ 6,\ 12)$ exists and is unique. Naturally the unique minimal $(1, 5;\ 10)$ design is the two-point deletion of $S(5,\ 6,\ 12)$.

2. Block sizes and upper bound on $g(1,\ 5;\ 10)$

Obviously for a minimal $(\lambda, \mu; v)$ design we consider blocks of size at least μ. We first state a useful result due to Woodall [2].

THEOREM A (Woodall). If $(1, \mu; v)$ design contains a block of length k then it contains at least $1 + h_1\ (\mu, k,\ v)$ blocks where

$$h_1(\mu,\ k,\ v) = (v-k)\binom{k}{\mu-1}\left\{1 - \frac{v-k-1}{2(k-\mu+2)}\right\}.$$

It is known [1] that there is a (unique) Steiner system $S(5,\ 6,\ 12)$. Deleting two points form $S(5,\ 6,\ 12)$ we get a $(1,\ 5;\ 10)$ design with 102 blocks. Thus $g(1,\ 5;$

10) $\leqslant 102$. Using this information along with theorem A we get the following:

PROPOSITION 2.1. A minimal (1, 5; 10) design cannot contain a block of size greater than 7.

3. The blocks of size 7 in a minimal (1, 5; 10) design

The following result is an immediate consequence of the definition of a (1, 5; 10) design.

PROPOSITION 3.1. There can be at most 3 blocks of size 7 in a (1, 5; 10) design.

We next successively rule out the existence of blocks of size 7 in a minimal (1, 5; 10) design. We use the following notation: n_i denotes the number of blocks of size i in a minimal $(\lambda, \mu; \nu)$ design, $\mu \leqslant i \leqslant \nu - 1$. Clearly in a minimal (1, 5; 10) design, where now the admissible block sizes are 5, 6 and 7, we must have

$$n_5 + \binom{6}{5} n_6 + \binom{7}{5} n_7 = \binom{10}{5} = 252. \tag{1}$$

PROPOSITION 3.2. In a minimal (1, 5; 10) design there cannot be 3 blocks of size 7.

Proof. Let, without loss of generality, the three blocks of size 7 be (1, 2, 3; 4, 5, 6, 7), (1, 2, 3, 4, 8, 9, 10) and (4, 5, 6, 7, 8, 9, 10). Now a block of size 6 is of the form (0, 0, 0, 8, 9, 10) or of the form $(0, 0, 0, 0, x, x)$ where $0 \in \{1, 2, 3, 4, 5, 6, 7\}$ and $x \in \{8, 9, 10\}$. It can be directly seen that a block of size 6 cannot be of the form (0, 0, 0, 8, 9, 10). It can also be directly seen that $n_6 \leqslant 3 \times 3 = 9$. Now, using (1) we see that with $n_7 = 3$, $n_6 \leqslant 5$ we get $n_5 \geqslant 135$. Since $g(1, 5; 10) \leqslant 102$, the result follows.

PROPOSITION 3.3. In a minimal (1, 5; 10) design there cannot be two blocks of size 7.

Proof. As before, without loss of generality, let the two blocks of size 7 be (1, 2, 3, 4, 5, 6, 7) and (1, 2, 3, 4, 8, 9, 10). Let m (respectively t) be the number of blocks of size 6 of the form (0, 0, 0, 8, 9, 10) (respectively $(0, 0, 0, 0, x, x)$ where $0 \in \{1, 2, 3, 4, 5, 6, 7\}$ and $x \in \{8, 9, 10\}$. As before it can be seen that $m \leqslant 3$ and $t \leqslant 3 \times 6 = 18$. So $n_6 \leqslant 21$. However, from (1) and the fact that $g(1, 5; 10) \leqslant 102$, we must have $n_6 \geqslant 22$, a contradiction.

PROPOSITION 3.4. In a minimal (1, 5; 10) design there cannot be exactly 1 block of size 7.

Proof. Suppose there is a unique block of size 7. Take it as (1, 2, 3, 4, 5, 6, 7). We have the following picture of blocks of size 6 and size 5.

l	m	a	b	c
0	0	0	0	0
0	0	0	0	0
0	0	x	0	0
x	0	x	x	0
x	x	x	x	x
x	x		n_5	
n_6				

Here l, m, a, b, c denote the number of blocks of respective type. We call any block among these l blocks an L-type block. M-type block is similarly defined. Note that 5-tuples not covered by the unique block of size 7 are of three types: $(0, 0, x, x, x)$, $(0, 0, 0, x, x)$ and $(0, 0, 0, 0, x)$. Counting them in two ways we get respectively,

(i) $3l + a = \binom{7}{2} = 21,$

(ii) $3l + 4m + b = \binom{7}{3} \times \binom{3}{2} = 105,$

(iii) $2m + c = \binom{7}{4} \times \binom{3}{1} = 105.$

Thus

$$6l + 6m + a + b + c = 231 = 6n_6 + n_5. \tag{2}$$

Observe that a given 3-tuple from 7 points is contained in four 4-tuples on those 7 points. We also note that if (i, j, k, x, x, x) is one of the L-type blocks then the triple i, j, k cannot be a part of any M-type block. To be precise, if (i, j, k, x, x, x) is a block then it rules out blocks (i, j, k, t, x, x) for any $t \in \{1, 2, \ldots, 7\} - \{i, j, k\}$. Thus

$$m \leqslant \binom{7}{4} - 4l = 35 - 4l. \tag{3}$$

Let $D_k(n)$ denote the maximum number of k-subsets of an n-set such that no $(k-1)$-tuple is repeated. It is easy to see that $D_4(7) = 7$. Thus

$$m \leqslant D_4(7) \times 3 = 21. \tag{4}$$

Now, let $m = 35 - 4l - i$, $0 \leqslant i \leqslant 35 - 4l$. The total number of blocks is

$$1 + l + m + a + b + c = 1 + l + (35 - 4l - i) + 231 - 6l - 6(35 - 4l - i), \text{using (2)}$$

$$= 57 + 15l + 5i. \tag{5}$$

If $l \geqslant 4$ then from (5) we see that the total number of blocks exceeds 102. Next, for $l \leqslant 3$ take $m = 21 - i$ in view of (4). The total number of blocks is then

$$1 + l + m + a + b + c = 1 + l + (21 - i) + 231 - 6l - 6(21 - i)$$

$$= 127 - 5l + 5i. \tag{6}$$

This rules out $l \leqslant 3$ for a minimal $(1, 5; 10)$ design. The proof of Proposition 3.4 is now complete.

Combining the above three results we get

THEOREM 3.5. In a minimal (1, 5; 10) design the block sizes are 6 and 5.

4. Minimal (1, 5; 10) designs with block sizes 6 and 5

We now have

$$6n_6 + n_5 = \binom{10}{5} = 252. \tag{7}$$

Since $n_5 + n_6 \leqslant 102$ we get

$$n_6 \geqslant 30. \tag{8}$$

Now let $(1, 2, 3, 4, 5, 6) = B_0$ be a block of size 6. The other blocks of size 6 fall into three types L, M and T as picturised below:

l	m	t
0	0	0
0	0	0
7	0	0
8	x	0
9	x	x
10	x	x

where l, m, t indicate the number of blocks of type L, M and T respectively, $0 \in \{1, 2, 3, 4, 5, 6\}$ and $x \in \{7, 8, 9, 10\}$. Clearly

$$l \leqslant 3. \tag{9}$$

Consider the following finer picture for the blocks of type M and type T.

m_1	m_2	m_3	m_4	t_1	t_2	t_3	t_4	t_5	t_6
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
7	7	7	8	0	0	0	0	0	0
8	8	9	9	7	7	7	8	8	9
9	10	10	10	8	9	10	9	10	10

It is easy to see that $D_3(6) = 4$. Hence we have

$$0 \leqslant m_i \leqslant 4 \ \forall i, \ 1 \leqslant i \leqslant 4. \tag{10}$$

It is easy to see that $D_4(6) = 3$, and therefore

$$0 \leqslant t_i \leqslant 3 \ \forall i, \ 1 \leqslant i \leqslant 6. \tag{11}$$

It can easily be seen that

$$
\begin{aligned}
m_1 = 4 &\Rightarrow t_1+t_2+t_4 \leqslant 3,\\
m_1 = 3 &\Rightarrow t_1+t_2+t_4 \leqslant 6,\\
m_2 = 4 &\Rightarrow t_1+t_3+t_5 \leqslant 3,\\
m_2 = 3 &\Rightarrow t_1+t_3+t_5 \leqslant 6,\\
m_3 = 4 &\Rightarrow t_2+t_3+t_6 \leqslant 3,\\
m_3 = 3 &\Rightarrow t_2+t_3+t_6 \leqslant 6,\\
m_4 = 4 &\Rightarrow t_4+t_5+t_6 \leqslant 3,\\
m_4 = 3 &\Rightarrow t_4+t_5+t_6 \leqslant 6.
\end{aligned}
\tag{12}
$$

We are now in a position to prove the following result which leads to the unique minimal (1, 5; 10) design.

THEOREM 4.1. $m+t \leqslant 26$. Further, the equality holds if and only if $m=8$ and $t=18$ with $m_i=2$, $t_j=3$, $\forall_i$, $1 \leqslant i \leqslant 4$, and for all j, $1 \leqslant j \leqslant 6$.

Proof. We consider two cases.

Case 1. Suppose some $m_i=4$.

If at least two m_i's are 4, then it follows from (11) and (12) that $t=\sum_{i=1}^{6} t_i \leqslant 3+3+3 = 9$ and then $m+t \leqslant 16+9=25$.

If precisely one m_i is 4 then $m \leqslant 4+3+3+3=13$ and $t=\sum_{i=1}^{6} t_i \leqslant 3+3\times 3=12$ giving $m+t \leqslant 25$.

Case 2. No m_i is 4.

If all the m_i's are 3, then $m=12$ and $2\sum_{i=1}^{6} t_i \leqslant 24$ giving $t \leqslant 12$ and hence $m+t \leqslant 24$.

If precisely three m_i's are 3, let without loss of generality, $m_1=m_2=m_3=3$. Then $m \leqslant 3+3+3+2=11$ and $2(t_1+t_2+t_3)+(t_4+t_5+t_6) \leqslant 6+6+6=18$. If $t=(t_1+t_2+t_3)+(t_4+t_5+t_6) \geqslant 15$, we get $t_1+t_2+t_3 \leqslant 3$. This implies that $t_4+t_5+t_6 \geqslant 12$, a contradiction to (10). Thus $t \leqslant 14$ and hence $m+t \leqslant 11+14=25$.

If the number of m_i's which equal 3 is one or two, then $m \leqslant 3+3+2+2=10$ and $t \leqslant 6+3\times 3=15$ giving $m+t \leqslant 25$.

If no m_i is 3, then $m \leqslant 8$ and $t \leqslant 18$ giving $m+t \leqslant 26$.

It is clear from the above proof that $m+t=26$ if and only if $m=8$ and $t=18$ with each $m_i=2$ and each $t_j=3$, $1 \leqslant i \leqslant 4$, $1 \leqslant j \leqslant 6$.

We now give our final result of this section.

THEOREM 4.2. In a minimal $g(1, 5; 10)$ design $n_6=1+l+m+t=1+3+8+18=30$ and $n_5=72$. Consequently $g(1, 5; 10)=102$.

Proof. From theorem 3.5 we know that a minimal (1, 5; 10) design has blocks of sizes 6 and 5 only. From theorem 4.1 we know that $n_6 = 1 + l + (m+t) \leqslant 1 + 3 + 26 = 30$. From (8) we have $n_6 \geqslant 30$. Thus $n_6 = 30$ (and $l = 3$). Consequently from (7) we get $n_5 = 72$ and it now follows that $g(1, 5; 10) = n_6 + n_5 = 30 + 72 = 102$.

5. Construction of the unique minimal (1, 5; 10) design

Start with the block $B_0 = (1, 2, 3, 4, 5, 6)$. We know that $l = 3$, $m = 8$, $t = 18$ with $m_i = 2\ \forall i$, $t_j = 3\ \forall j, 1 \leqslant l \leqslant 4, 1 \leqslant j \leqslant 6$. Therefore, without loss of generality, we have the following partial set-up for the $n_6 = 30$ blocks of size 6 in a minimal (1, 5; 10) design. We have written blocks as columns.

B_0							m				
1	1	3	5	1	2	1	2	1	2	1	2
2	2	4	6	3	4	3	4	4	3	4	3
3	7	7	7	5	6	6	5	5	6	6	5
4	8	8	8	7	7	7	7	7	7	8	8
5	9	9	9	8	8	8	8	9	9	9	9
6	10	10	10	9	9	10	10	10	10	10	10

(13)

								t									
0																	
0																	
0																	
0																	
7	7	7	7	7	7	7	7	7	8	8	8	8	8	8	9	9	9
8	8	8	9	9	9	10	10	10	9	9	9	10	10	10	10	10	10

(14)

where the blank spaces under t are to be filled with suitable choices of 4-tuples of 0's, $0 \in \{1, 2, 3, 4, 5, 6\}$. Remembering all the time that no 5-tuple is to be repeated, it can directly be seen that the blank spaces under t can be filled in a unique way. Without loss of generality, the unique completion of the blanks under t which give us blocks of type T is as follows:

1	1	2	1	1	2	3	1	1	3	1	1	1	1	2	1	2	1
2	4	3	2	3	3	4	2	2	4	2	2	2	3	3	2	4	3
3	5	5	5	4	4	5	3	4	5	3	4	5	4	4	3	5	5
4	6	6	6	6	5	6	5	6	6	6	5	6	5	6	4	6	6
7	7	7	7	7	7	7	7	7	8	8	8	8	8	8	9	9	9
8	8	8	9	9	9	10	10	10	9	9	9	10	10	10	10	10	10

The $n_6 = 30$ blocks given by (12) and (14) together completely describe the

unique minimal (1, 5; 10) design. The $n_5 = 72$ blocks of size 5 are precisely those 5-tuples which are not covered by the 30 blocks of size 6.

Acknowledgement

This research was supported by a fellowship from UGC, New Delhi.

References

[1] Lüneberg H, Über die Gruppen Von Mathieu, *J. Algebra* **10** (1968) 194–210
[2] Woodall D R, The λ-μ problem, *J. London Math. Soc.* (2) **1** (1969) 505–519

Proc. Indian Acad. Sci. (Math. Sci.), Vol. 95, No. 1, September 1986, pp. 53–59.

Some generating functions of Jacobi polynomials

BIDYUT GUHA THAKURTA

Department of Mathematics, RKMVC College, Rahara, 24 Parganas, West Bengal, India.

MS received 30 September 1985

Abstract. In this paper, Weisner's group-theoretic method of obtaining generating functions is utilized in the study of Jacobi polynomials $P_n^{(\alpha,\beta)}(x)$ by giving suitable interpretations to the index (n) and the parameter (β) to find out the elements for constructing a six-dimensional Lie algebra.

Keywords. Jacobi polynomials; generating function.

1. Introduction

It is known that the Jacobi polynomials, defined by [9],

$$P_n^{(\alpha,\beta)}(x) = \frac{(1+\alpha)_n}{n!}\ {}_2F_1\left[\begin{matrix} -n,\ 1+\alpha+\beta+n; & \dfrac{1-x}{2} \\ 1+\alpha; & \end{matrix}\right] \tag{1}$$

satisfies the following ordinary differential equation:

$$(1-x^2)\frac{\mathrm{d}^2u}{\mathrm{d}x^2} + \{\beta-\alpha-(2+\alpha+\beta)x\}\frac{\mathrm{d}u}{\mathrm{d}x} + n(1+\alpha+\beta+n)u = 0. \tag{2}$$

The object of the present note is to derive some generating functions, which are believed to be new, of Jacobi polynomials as defined in (1) by suitably interpreting the index (n) and the parameter (β) simultaneously with the help of Weisner's group-theoretic method [10] (for previous work on $P_n^{(\alpha,\beta)}(x)$ by the same method see [1, 2, 4–8]). The main result (obtained by finding a set of infinitesimal operators A_{ij} ($i=1, 2; j=1, 2, 3$) constituting a Lie algebra) of our investigation is given in §3.

2. Group-theoretic method

Replacing $\mathrm{d}/\mathrm{d}x$ by $\partial/\partial x$, β by $y(\partial/\partial y)$, n by $z(\partial/\partial z)$ and $u(x,y,z)$ by $v(x,y,z)$ in (1) we get the following partial differential equation:

$$(1-x^2)\frac{\partial^2 v}{\partial x^2} + y(1-x)\frac{\partial^2 v}{\partial y\,\partial x} + yz\frac{\partial^2 v}{\partial z\,\partial y} + z^2\frac{\partial^2 v}{\partial z^2}$$

$$-\{(1+x)\alpha+2x\}\frac{\partial v}{\partial x} + (\alpha+2)z\frac{\partial v}{\partial z} = 0. \tag{3}$$

Thus $v_1(x, y, z) = P_n^{(\alpha,\beta)}(x)y^{\beta}z^n$ is a solution of the differential equation (3) since $P_n^{(\alpha,\beta)}(x)$ is a solution of (2).

Now by using the following differential recurrence relations [1, 9]:

$$\frac{d}{dx} P_n^{(\alpha,\beta)}(x) = \frac{1}{x-1}[nP_n^{(\alpha,\beta)}(x) - (\alpha+n)P_{n-1}^{(\alpha,\beta+1)}(x)], \tag{4}$$

$$\frac{d}{dx} P_n^{(\alpha,\beta)}(x) = \frac{1}{1-x^2}[\{(1+\alpha+\beta+n)(x+1)-2\beta\} \times P_n^{(\alpha,\beta)}(x) - 2(n+1)P_{n+1}^{(\alpha,\beta-1)}(x)], \tag{5}$$

$$\frac{d}{dx} P_n^{(\alpha,\beta)}(x) = \frac{1}{1-x}[(1+\alpha+\beta+n)\{P_n^{(\alpha,\beta)}(x) - P_n^{(\alpha,\beta+1)}(x)], \tag{6}$$

$$\frac{d}{dx} P_n^{(\alpha,\beta)}(x) = \frac{1}{1-x^2}[-\{n(x+1)+2\beta\} \times P_n^{(\alpha,\beta)}(x) + 2(\beta+n)P_n^{(\alpha,\beta-1)}(x)]. \tag{7}$$

We define the infinitesimal operators $A_{ij}(i=1, 2;\ j=1, 2, 3)$ as follows:

$$\begin{aligned}
A_{11} &= y(\partial/\partial y),\\
A_{12} &= (x-1)y\, z^{-1}(\partial/\partial x) - y(\partial/\partial z)\\
A_{22} &= (1-x^2)y^{-1}z(\partial/\partial x) - z(x-1)(\partial/\partial y)\\
&\qquad - (1+x)y^{-1}z^2(\partial/\partial z) - (1+\alpha)(1+x)y^{-1}z\\
A_{21} &= z(\partial/\partial z)\\
A_{13} &= (1-x^2)y^{-1}(\partial/\partial x) + 2(\partial/\partial y) + (1+x)y^{-1}z(\partial/\partial z),\\
A_{23} &= (1-x)y(\partial/\partial x) - y^2(\partial/\partial y) - yz(\partial/\partial z) - (1+\alpha)y,
\end{aligned}$$

such that

$$\begin{aligned}
A_{11}(P_n^{(\alpha,\beta)}(x)\, y^{\beta} z^n) &= \beta P_n^{(\alpha,\beta)}(x)\, y^{\beta} z^n,\\
A_{12}(P_n^{(\alpha,\beta)}(x)\, y^{\beta} z^n) &= -(\alpha+n)P_{n-1}^{(\alpha,\beta+1)}(x)\, y^{\beta+1} z^{n-1},\\
A_{22}(P_n^{(\alpha,\beta)}(x)\, y^{\beta} z^n) &= -2(n+1)P_{n+1}^{(\alpha,\beta-1)}(x)\, y^{\beta-1} z^{n+1},\\
A_{21}(P_n^{(\alpha,\beta)}(x)\, y^{\beta} z^n) &= n\, P_n^{(\alpha,\beta)}(x)\, y^{\beta} z^n,\\
A_{13}(P_n^{(\alpha,\beta)}(x)\, y^{\beta} z^n) &= 2(\beta+n)P_n^{(\alpha,\beta-1)}(x)\, y^{\beta-1} z^n,\\
A_{23}(P_n^{(\alpha,\beta)}(x)\, y^{\beta} z^n) &= -(1+\alpha+\beta+n)\, P_n^{(\alpha,\beta+1)}(x)\, y^{\beta+1} z^n.
\end{aligned}$$

Now we shall find the commutator relations. Using the notation:

$$[A, B]u = (AB - BA)u,$$

we have

$$\begin{aligned}
&[A_{11}, A_{12}] = A_{12}, && [A_{12}, A_{23}] = 0,\\
&[A_{11}, A_{13}] = -A_{13}, && [A_{21}, A_{22}] = A_{22},\\
&[A_{11}, A_{21}] = 0, && [A_{21}, A_{23}] = 0,\\
&[A_{11}, A_{22}] = -A_{22}, && [A_{13}, A_{21}] = 0,\\
&[A_{11}, A_{23}] = A_{23}, && [A_{13}, A_{22}] = 0,\\
&[A_{12}, A_{13}] = 0, && [A_{13}, A_{23}] = -2(2A_{11}+2A_{21}+1+\alpha),\\
&[A_{12}, A_{21}] = A_{12}, && [A_{22}, A_{23}] = 0.\\
&[A_{12}, A_{22}] = 2(2A_{21}+(1+\alpha)), &&
\end{aligned} \tag{8}$$

From the above commutator relations we state the following theorem:

THEOREM

The set of operators $\{1, A_{ij}\ (i = 1, 2;\ j = 1, 2, 3)\}$ where 1 stands for the identity operator, generates a Lie algebra.

It can be easily shown that the partial differential operator L, given by

$$L = (1-x^2)\frac{\partial^2}{\partial x^2} + y(1-x)\frac{\partial^2}{\partial y\partial x} + yz\frac{\partial^2}{\partial z\partial y} + z^2\frac{\partial^2}{\partial z^2}$$

$$-\left\{(1+x)\,\alpha + 2x\right\}\frac{\partial}{\partial x} + (\alpha+2)\,z\frac{\partial}{\partial z}$$

can be expressed as follows:

$$(x-1)\,Lu = (A_{22}A_{12} - 2A_{21}^2 - 2\alpha\,A_{21})u$$

and $$(x-1)\,Lu = (-A_{13}A_{23} - 2(A_{11}+A_{21}+1)(A_{11}+A_{21}+1+\alpha))u, \tag{9}$$

From (8) and (9), one can easily verify that the operators A_{ij} $(i=1, 2; j=1, 2, 3)$ commute with $(x-1)L$, i.e.,

$$\left[(x-1)L, A_{ij}\right] = 0, \tag{10}$$

The extended form of the groups generated by A_{ij} $(i=1, 2; j=1, 2, 3)$ are given by

$$\exp(a_{11}A_{11})\ u(x, y, z) = u(x, \exp(a_{11})\ y, z), \tag{11}$$

$$\exp(a_{21}A_{21})\ u(x, y, z) = u(x, y, \exp(a_{21})\ z), \tag{12}$$

$$\exp(a_{12}A_{12})\ u(x, y, z) = u\left(\frac{zx-a_{12}\,y}{z-a_{12}\,y}, y, z-a_{12}\,y\right), \tag{13}$$

$$\exp(a_{22}A_{22})\ u(x,\ y,\ z)=u\left(\frac{y}{y+a_{22}\ (1+x)z}\right)^{\alpha+1} u\left(\frac{xy+a_{22}\ (1+x)z}{y+a_{22}\ (1+x)z},\right.$$

$$\left.\frac{y(y+2a_{22}z)}{y+a_{22}\ (1+x)z},\quad \frac{yz}{y+a_{22}\ (1+x)z}\right), \tag{14}$$

$$\exp(a_{13}A_{13})\ u(x,\ y,\ z)\ =\ u\left(\frac{xy+a_{13}(1+x)}{y+a_{13}(1+x)},\ y+2a_{13},\right.$$

$$\left.\frac{z\{y+a_{13}(1+x)\}}{y}\right) \tag{15}$$

$$\exp(a_{23}A_{23})\ u(x,\ y,\ z)\ =\ (1+a_{23}y)^{-\alpha-1}$$

$$\times u\left(\frac{x+a_{23}y}{1+a_{23}y},\ \frac{y}{1+a_{23}y},\ \frac{z}{1+a_{23}y}\right) \tag{16}$$

where $a_{ij}(i = 1, 2;\ j = 1, 2, 3)$ are arbitrary constants. Thus we have

$$\exp(a_{23}A_{23})\ \exp(a_{13}A_{13})\ \exp(a_{22}A_{22})\ \exp(a_{12}A_{12})\ \exp(a_{21}A_{21})$$
$$\times \exp(a_{11}A_{11})\ u(x,\ y,\ z)\ =\ [y/\{y(1+a_{23}y)+a_{22}z(1+x+2a_{23}y)\}\ ]^{\alpha+1}$$
$$\times u(\xi,\ \eta,\ \zeta) \tag{17}$$

where

$$\xi = \frac{\begin{array}{c} z[y\{y(x+a_{23}y)+a_{13}(1+a_{23}y)\,(1+x+2a_{23}y)\}+a_{22}z(1+x+2a_{23}y) \\ \{y+a_{13}(1+x+2a_{23}y)\}]-a_{12}[y\{y+2a_{13}(1+a_{23}y)\} \\ +2a_{22}z\{y+a_{13}(1+x+2a_{23}y)\}][y(1+a_{23}y)+a_{22}z(1+x2a_{23}y)] \end{array}}{\begin{array}{c} [y(1+a_{23}y)+a_{22}z(1+x+2a_{23}y)\,][z\{y+a_{13}(1+x+2a_{23}y)\} \\ -a_{12}\{y(\,y+a_{13}(1+a_{23}y)\,)+2a_{22}z(\,y+a_{13}(1+x+2a_{23}y)\,)\}], \end{array}}$$

$$\eta = \exp(a_{11})\,\frac{y\{y+2a_{13}(1+a_{23}y)\}+2a_{22}z\{y+a_{13}(1+x+2a_{23}y)\}}{y(1+a_{23})+a_{22}z(1+x+2a_{23}y)} \tag{18}$$

and

$$\zeta = \exp(a_{21})\,\frac{\begin{array}{c} z\{y+a_{13}(1+x+2a_{23}y)\}-a_{12}[y\{y+2a_{13}(1+a_{23}y)\} \\ +2a_{22}z\{y+a_{13}(1+x+2a_{23}y)\}] \end{array}}{y(1+a_{23}y)+a_{22}z(1+x+2a_{23}y)} \tag{19}$$

3. Generating functions

From (3) $v(x,\ y,\ z)=P_n^{(\alpha,\ \beta)}(x)y^{\beta}z^n$ is a solution of the system:

$$\begin{array}{lll} Lv=0; & Lv=0; & Lv=0; \\ (A_{11}-\beta)v=0; & (A_{21}-n)v=0; & (A_{11}+A_{21}-\beta-n)v=0. \end{array}$$

From (13) we easily get

$$[(x-1)L]\, S(P_n^{(\alpha,\beta)}(x) y^{\beta} z^n) = S[(x-1)L]\,(P_n^{(\alpha,\beta)}(x)\, y^{\beta} z^n) = 0,$$

where

$$S = \exp(a_{23}A_{23}) \exp(a_{13}A_{13}) \exp(a_{22}A_{22}) \exp(a_{12}A_{12}) \exp(a_{21}) \exp(a_{11}).$$

Therefore the transformation $S(P_n^{(\alpha,\beta)}(x) y^{\beta} z^n)$ is also annulled by L.

Now by putting $a_{11}=a_{21}=0$ and replacing $u(x, y, z)$ by $P_n^{(\alpha,\beta)}(x) y^{\beta} z^n$ in (19) we get

$$\begin{aligned}
&\exp(a_{23}A_{23}) \exp(a_{13}A_{13}) \exp(a_{22}A_{22}) \exp(a_{12}A_{12})(P_n^{(\alpha,\beta)}(x)\, y^{\beta} z^n)\\
&= y^{\alpha+1}[y(1+a_{23}y)+a_{22}z(1+x+2a_{23}y)]^{-1-\alpha-\beta-n}\\
&\quad\times[y\{y+2a_{13}(1+a_{23}y)\}+2a_{22}z\{y+a_{13}(1+x+2a_{23}y)\}]^{\beta}\\
&\quad\times(z\{y+a_{13}(1+x+2a_{23}y)\}-a_{12}[y\{y+2a_{13}(1+a_{23}y)\}\\
&\quad+2a_{22}z\{y+a_{13}(1+x+2a_{23}y)\}])^n\, P_n^{(\alpha,\beta)}(\xi).
\end{aligned} \tag{20}$$

But,

$$\begin{aligned}
&\exp(a_{23}A_{23}) \exp(a_{13}A_{13}) \exp(a_{22}A_{22}) \exp(a_{12}A_{12})\,(P_n^{(\alpha,\beta)}(x) y^{\beta} z^n)\\
&= \sum_{s=0}^{\infty} \frac{a_{23}^s(-1)^s}{s!}(1+\alpha+\beta-r+n)_s \sum_{r=0}^{\infty} \frac{a_{13}^r}{r!}\, 2^r(n+\beta-r+1)_r\\
&\times \sum_{m=0}^{\infty} \frac{a_{22}^m}{m!}(-2)^m (n-k+1)_m \sum_{k=0}^{n+m} \frac{a_{12}^k}{k!}(-\alpha-n)_k\\
&\times P_{n-k+m}^{(\alpha,\beta+k-m-r+s)}(x)\, y^{\beta+k-m-r+s} z^{n-k+m}
\end{aligned} \tag{21}$$

Equating (20) and (21) we get

$$\begin{aligned}
&[y/\{y(1+a_{23}y)+a_{22}z(1+x+2a_{23}y)\}]^{1+\alpha+\beta+n}\Big[y+2a_{13}(1+a_{23}y)+2a_{22}\frac{z}{y}\\
&\times\{y+a_{13}(1+x+2a_{23}y)\}\Big]^{\beta}\Big(\frac{z}{y}\{y+a_{13}(1+x+2a_{23}y)\}\\
&-a_{12}[y+2a_{13}(1+a_{23}y)\\
&+2a_{22}\frac{z}{y}\{y+a_{13}(1+x+2a_{23}y)\}]\Big)^n P_n^{(\alpha,\beta)}(\xi)\\
&= \sum_{s=0}^{\infty} \frac{(-a_{23})^s}{s!}(1+\alpha+\beta-r+n)_s \sum_{r=0}^{\infty} \frac{(2a_{13})^r}{r!}(n+\beta-r+1)_r\\
&\times \sum_{m=0}^{\infty} \frac{(-2a_{22})^m}{m!}(n-k+1)_m \sum_{k=0}^{n+m} \frac{(a_{12})^k}{k!}(-\alpha-n)_k\\
&\times P_{n-k+m}^{(\alpha,\beta+k-m-r+s)}(x)\, y^{\beta+k-m-r+s} z^{n-k+m},
\end{aligned} \tag{22}$$

where

$$\xi = \frac{\begin{array}{l} z[y\{y(x+a_{23}y)+a_{13}(1+a_{23}y)(1+x+2a_{23}y)\}+a_{22}z(1+x+2a_{23}y) \\ \{y+a_{13}(1+x+2a_{23}y)\}]-a_{12}[y\{y+2a_{13}(1+a_{23}y)\} \\ +2a_{22}z\{y+a_{13}(1+x+2a_{23}y)\}]\ [y(1+a_{23}y)+a_{22}z(1+x+2a_{23}y)] \end{array}}{\begin{array}{l} [y(1+a_{23}y)+a_{22}z(1+x+2a_{23}y)][z\{y+a_{13}(1+x+2a_{23}y)\} \\ -a_{12}\{y(y+2a_{13}(1+a_{23}y))+2a_{22}z(y+a_{13}(1+x+2a_{23}y))\}] \end{array}}$$

The above generating function does not appear before from which we can get a large number of particular generating relations by attributing different values to a_{12}, a_{22}, a_{13}, a_{23}.

4. Derivation of some known results

Case 1.

Putting $a_{13}=a_{23}=0$ in (22), we get

$$\left(\frac{y}{y+a_{22}z(1+x)}\right)^{1+\alpha+n+\beta} (y+2a_{22}z)^{\beta}\{z-a_{12}(y+2a_{22}z)\}^{n}$$
$$\times P_n^{(\alpha,\beta)}\left(\frac{z\{xy+a_{22}z(1+x)\}-a_{12}(y+2a_{22}z)\{y+a_{22}z(1+x)\}}{\{y+a_{22}z(1+x)\}\{z-a_{12}(y+2a_{22}z)\}}\right)$$
$$=\sum_{m=0}^{\infty}\frac{(a_{22})^m}{m!}(-2)^m(n-k+1)_m\sum_{k=0}^{n+m}\frac{(a_{12})^k}{k!}(-\alpha-n)_k$$
$$\times P_{n-k+m}^{(\alpha,\beta+k-m)}(x)\,y^{\beta+k-m}\,z^{n-k+m} \tag{23}$$

which is the main result obtained by Chakraborti [1].

Case 2.

Putting $a_{12}=a_{13}=a_{23}=0$, $a_{22}=1$ and writing $z/y=-u/2$ we get

$$\frac{(1-u)^{\beta}}{\{1-\frac{1}{2}u(1+x)\}^{1+\alpha+\beta+n}}\,P_n^{(\alpha,\beta)}\left(\frac{x-\frac{1}{2}u(1+x)}{1-\frac{1}{2}u(1+x)}\right)$$
$$=\sum_{m=0}^{\infty}\frac{1}{m!}(n+1)_m\,P_{n+m}^{(\alpha,\beta-m)}(x)u^m.$$

Finally putting $n=0$ we get

$$(1-u)^{\beta}(1-\tfrac{1}{2}u(1+x))^{-1-\alpha-\beta}=\sum_{m=0}^{\infty} P_m^{(\alpha,\,\beta-m)}(x)u^m, \tag{24}$$

which is a well-known formula due to Feldhim [3].

Acknowledgements

I am greatly indebted to Dr A K Chongdar, Department of Mathematics, Bangabasi Evening College, Calcutta for introducing me to this subject and for suggesting this particular area of investigation as well as for his constant encouragement and guidance.

References

[1] Chakraborti A B, group theoretic study of certain generating functions of Jacobi polynomials, *J. Indian Inst. Sci.* **B64** (1983) 97–103

[2] Chongdar A K, Group theoretic study for certain generating functions, *Bull. Calcutta Math. Soc.* **77** (1985) 151–157

[3] Feldhim E, Relations entre les polynomes de Jacobi, Laguerre et Hermite, *Acta Math.* **75** (1943) 117–198

[4] Ghosh B, Group theoretic study of certain generating functions of Jacobi polynomials, *Bull. Calcutta Math. Soc.* **75** (1983) 227–231

[5] Ghosh B, On a class of trilateral generating functions for Jacobi polynomials, *Ranchi Univ. Math. J.* **14** (1983) 11–16

[6] Ghosh B, Group theoretic origins of certain generating functions of Jacobi polynomials I, *Pure Math. Manuscript* **3** (1984) 139–144

[7] Ghosh B, Group theoretic origins of certain generating functions of Jacobi polynomials, II, (1986) communicated

[8] Ghosh B, Group theoretic origins of certain generating functions of Jacobi polynomials III, *Pure Math. manuscript* **3** (1984) (to appear)

[9] Rainville E D, *Special functions*, Chelsea Publishing Co. Brons, New York (1960)

[10] Weisner L, Group theoretic origin of certain generating functions, *Pacific J. Math.* **5** (1955) 1033–1039

Proc. Indian Acad. Sci. (Math. Sci.), Vol. 95, No. 1, September 1986, pp. 61–77.

Cohomology of the moduli of parabolic vector bundles

NITIN NITSURE

School of Mathematics, Tata Institute of Fundamental Research, Homi Bhabha Road, Bombay 400 005, India

MS received 25 March 1985

Abstract. The purpose of this paper is to compute the Betti numbers of the moduli space of *parabolic vector bundles* on a curve (see Seshadri [7], [8] and Mehta & Seshadri [4]), in the case where every semi-stable parabolic bundle is necessarily stable. We do this by generalizing the method of Atiyah and Bott [1] in the case of moduli of ordinary vector bundles. Recall that (see Seshadri [7]) the underlying topological space of the moduli of parabolic vector bundles is the space of equivalence classes of certain unitary representations of a discrete subgroup Γ which is a lattice in PSL (2, **R**). (The lattice Γ need not necessarily be co-compact).

While the structure of the proof is essentially the same as that of Atiyah and Bott, there are some difficulties of a technical nature in the parabolic case. For instance the Harder–Narasimhan stratification has to be further refined in order to get the connected strata. These connected strata turn out to have different codimensions even when they are part of the same Harder–Narasimhan strata.

If in addition to 'stable = semistable' the rank and degree are coprime, then the moduli space turns out to be torsion-free in its cohomology.

The arrangement of the paper is as follows. In § 1 we prove the necessary basic results about algebraic families of parabolic bundles. These are generalizations of the corresponding results proved by Shatz [9]. Following this, in § 2 we generalize the analytical part of the argument of Atiyah and Bott (§ 14 of [1]). Finally in § 3 we show how to obtain an inductive formula for the Betti numbers of the moduli space. We illustrate our method by computing explicitly the Betti numbers in the special case of rank = 2, and one parabolic point.

Keywords. Cohomology; parabolic vector bundles; moduli space; Betti numbers; algebraic family; Sobolev spaces.

1. Algebraic families of parabolic bundles

For the basic definitions and properties of vector bundles with parabolic structures ('parabolic bundles' for short) see Seshadri [7] and [8], and Mehta and Seshadri [4]. For simplicity we assume throughout that the parabolic structures of the parabolic bundles considered are over only a single point P of the curve X. All our arguments generalize trivially to the case where there are more parabolic points on X.

As in [8], each parabolic bundle E on X has a unique (parabolic) Harder–Narasimhan filtration $0 \subset E_1 \subset \ldots \subset E_r = E$ where E_i are parabolic subbundles such that (i) each quotient parabolic bundle E_i/E_{i-1} is semi-stable (ii) the inequality par $\mu(E_i/E_{i-1}) >$ par $\mu(E_{i+1}/E_i)$ holds for each i, where par μ = (par deg/rank)

Hence to each parabolic bundle E, we may associate a convex polygon HNP (E) in $\mathbf{R}^2$ as defined by Shatz [9]. The vertices of HNP (E) (= the Harder–Narasimhan polygon of E) are the points $(0,0)$, (rank E_1, par deg E_1), . . . , (rank E_r, par deg E_r). The polygons corresponding to parabolic bundles of a given rank and parabolic degree have a natural partial ordering as defined by Shatz [9], namely if λ_1 and λ_2 are two such polygons then $\lambda_1 \geq \lambda_2$ if all the vertices of λ_2 lie on or below the polygon λ_1. If $\lambda_1 \geq \lambda_2$, we say that λ_1 dominates λ_2. The polygon λ = HNP (E) is also called the *Harder–Narasimhan type* of the parabolic bundle E.

PROPOSITION 1.1. Let E be a parabolic bundle. Then any parabolic subbundle of E lies on or below the Harder–Narasimhan polygon of E. In particular, the polygon corresponding to any filtration of E by parabolic subbundles is dominated by the Harder–Narasimhan polygon HNP (E) of E.

Proof. If F_1 and F_2 are parabolic subbundles of E, and if the vector subbundles $F_1 \vee F_2$ and $F_1 \cap F_2$ of E (as defined in Langton [3]) are given the induced parabolic structure, then it is easy to see that par deg $(F_1 \vee F_2)$+par deg $(F_1 \cap F_2) \geq$ par deg (F_1)+par deg (F_2). Now the proof of proposition 2 follows from the proof of the theorem 2 of Shatz [9].

We want to study how the Harder–Narasimhan type changes within an algebraic family of parabolic bundles. For this, the following observation is basic.

Remark 1.2. Let E be a vector bundle on a scheme S and let F_1 and F_2 be subbundles. Let $\phi\colon F_1 \to E/F_2$ be the natural map. Then the function $s \mapsto \mathrm{rank}_{k(s)}\phi(s)$ is lower-semicontinuous on S. (Equivalently, the function $s \mapsto \dim_{k(s)} F_1(s) \cap F_2(s)$ is upper semicontinuous).

Using this, it is easy to prove the following:

PROPOSITION 1.3. Let E_T be a family of parabolic bundles parametrized by the scheme T which is the spectrum of a discrete valuation ring. Let ξ and ξ_0 be the generic and special points of T. Let G be a coherent torsion-free quotient sheaf of E_T on $X \times T$ which is flat over T. Let G_ξ and G_{ξ_0} be the restrictions of G to $X \times \xi$ and $X \times \xi_0$. Let G'_{ξ_0} be the vector bundle on $X \times \xi_0$ generically generated by G_{ξ_0}. Let G_ξ and G'_{ξ_0} be given the induced parabolic structures, as quotients of E_ξ and E_{ξ_0} respectively. Then par $\deg(G_\xi) \geq$ par $\deg(G'_{\xi_0})$.

PROPOSITION 1.4. Let E_T be a family of parabolic bundles on X parametrized by T, where T is the spectrum of a discrete valuation ring. Let ξ, ξ_0 be the generic and special points of T. Then $\mathrm{HNP}(E_{\xi_0}) \geq \mathrm{HNP}(E_\xi)$.

Proof. Let $0 \subset E_{\xi,1} \subset \ldots \subset E_{\xi,r} = E_\xi$ be the Harder–Narasimhan filtration over the generic point $\xi \in T$. Then using completeness of appropriate quot schemes it follows (Shatz [9], proposition 9) that there exists a filtration $0 \subset E_1 \subset \ldots \subset E$ of the bundle $E \to T \times X$ by subbundles (i.e. torsion free, coherent subsheaves

having a torsion free quotient) which induces the given filtration $0 \subset E_{\xi,1} \subset \ldots \subset E_{\xi,r} = E_\xi$ over the generic point. It follows from the proposition 1.3 that if $(E_i)'_{\xi_0}$ is the subbundle of E_0 generically generated by $(E_i)_{\xi_0}$, then par deg $(E_i)'_{\xi_0} \geq$ par deg $E_{\xi,i}$. Now the proposition follows from proposition 1.1.

Remark 1.5. Let k be an infinite field, let K/k be an extension field, let X be a curve over k, E a parabolic bundle on X, and E_K its pull back to X_K. Then it is proved by Seshadri [8] that the SCSS subbundle (i.e. β-subbundle) of the parabolic bundle E_K is the pull back of the SCSS subbundle of E. Hence the Harder–Narasimhan filtration of E_K is the pull back of that of E, and in particular E_K is semi-stable if and only if E is semi-stable.

From this remark and the proposition 1.4, we get the following result analogous to the theorem 3 of Shatz [9].

PROPOSITION 1.6. Let E_S be a family of parabolic bundles on X parametrized by a scheme S, and let s, $s_0 \in S$ such that s_0 lies in the closure of s. Then $\mathrm{HNP}(E_{s_0}) \geq \mathrm{HNP}(E_s)$.

Proof. There exists a morphism $T \to S$ of schemes where T is the spectrum of a *d.v.r.* sending the generic and special points ξ and ξ_0 respectively to s and s_0. Let E_T be the pulled back parabolic family. Then by Prop. 1.4 and remark 1.5,

$$\mathrm{HNP}(E_{s_0}) = \mathrm{HNP}(E_{\xi_0}) \geq \mathrm{HNP}(E_\xi) = \mathrm{HNP}(E_s).$$

PROPOSITION 1.7. Let E_S be a family of parabolic bundles on X parametrized by a scheme S. Then all points $s \in S$ such that E_s is semi-stable form an open subset of S.

Proof. Let $d = \deg(E_s)$, $r = \mathrm{rank}(E_s)$, which are constants. Then from the assumption that S is noetherian, the upper semi-continuity of dim $H^1(X_s, E_s)$, and the Riemann–Roch theorem it is easy to derive (see Narasimhan and Ramanathan [5]), that there exist an integer m such that for any coherent quotient sheaf E''_s of E_s for any $s \in S$ deg $(E''_s) \geq m$.

Now if a parabolic bundle E_s is not semi-stable, then there exists a quotient E''_s such that

$$\frac{\text{par deg }(E_s)}{\text{rank }(E_s)} \geq \frac{\text{par deg }(E''_s)}{\text{rank }(E''_s)}.$$

Now, par deg $(E''_s) \geq$ deg (E''_s). Hence we would have

$$\deg\ (E''_s) \leq \text{par deg } (E_s) \leq d+r \text{ where } r = \text{rank } E.$$

Hence in the family E_S, a parabolic bundle E_s is not semi-stable if and only if it has a quotient parabolic bundle E''_s with

(i) par $\mu(E) \geq$ par $\mu(E''_s)$,
(ii) $d+r \geq$ deg $(E''_s) \geq m$.

Let Q/S be the relative Quot scheme of quotients of the E_s of ranks $r_1 \leq r$, and degrees d_1 with $d+r \geq d_1 \geq m$. Then we know that Q is projective over S.

Let G be the universal sheaf over $X \times Q$. As G is coherent, and flat over Q, the points $t \in Q$ such that G_t is a torsion free sheaf on X_t form an open subset $Y \subseteq Q$ as follows from EGA IV_3, theorem 12.2.1 (iii) or (iv).

Let π: $Q \to S$ be the structure morphism. Then for each $y \in Y$, G_y is a quotient bundle of $(\pi^* E)_y$, and hence acquires a parabolic structure. Note that as G is a flat quotient, the function $y \mapsto \deg(G_y)$ is a locally constant function. It now easily follows from remark 1.2 that $y \mapsto \operatorname{par} \deg(G_y)$ is a lower semi-continuous function on Y. The function $y \mapsto \operatorname{rank}(G_y)$ is locally constant. Hence, the function $y \mapsto$ par $\mu(G_y)$ is also a lower semi-continuous function on Y. Hence

$$Z = \{y \in Y \mid \operatorname{par} \mu(E) \geq \operatorname{par} \mu(G_y)\} \text{ is a closed subset of } Y.$$

Note that E_s is unstable ($=$ non semi-stable) if and only if there exists some $y \in Z$ such that $\pi(y) = s$. Let S_{unstable} be the set of all $s \in S$ such that E_s is not semi-stable. Hence $\pi(Z) = S_{\text{unstable}}$. We wish to prove that S_{unstable} is closed in S. For this, let $\widetilde{Z}$ be the closure of Z in Q. As $\pi: Q \to S$ is projective, $\pi(\widetilde{Z})$ is closed in S. We claim that $\pi(\widetilde{Z}) \subseteq S_{\text{unstable}}$ which would complete the proof. So let $y_0 \in \widetilde{Z}$ Then as Z is a locally closed subscheme of Q, there exists some $y \in Z$ such that y_0 lies in the closure of y. Hence $\pi(y_0)$ lies in the closure of $\pi(y) \in S_{\text{unstable}}$. Hence by Prop. 1.6, $E_{\pi(y_0)}$ is unstable, which completes the proof.

Proposition 1.8. Let E_S be a family of parabolic bundles parametrized by an irreducible scheme S. Then there exists an open subset $U \subseteq S$ such that

(i) for all $s \in U$, $\mathrm{HNP}(E_s) =$ constant.

(ii) there exists an algebraic filtration on the restricted family E_U which specializes to the Harder–Narasimhan filtration at each point of U.

Proof. Let ξ be the generic point of S. Let $0 \subset E_{1,\xi} \subset \ldots \subset E_{r,\xi} = E_\xi$ be the Harder–Narasimhan filtration of the parabolic bundle E on X. As ξ is the generic point of S, there exists an open subset $V \subseteq S$, and a filtration $0 \subset E_1 \subset \ldots \subset E_r = E_V$ of E_V by vector subbundles which restricts to the Harder–Narasimhan filtration at the generic point ξ. Give each $E_{i,s}$, for $s \in V$, the induced parabolic structure. Then it is easy to see that there exists an open subset $V' \subseteq V$ such that $\operatorname{par}\deg(E_{i,s}) = \operatorname{par}\deg(E_{i,\xi})$ for all s in V'. Now, each quotient $E_{i,\xi}/E_{i-1,\xi}$ is semi-stable. Hence by proposition 1.7, there exists an open subset $U \subseteq V$ such that each $E_{i,s}/E_{i-1,s}$ is semi-stable for $s \in U$. Further, as $\mu(E_{i,\xi}/E_{i-1,\xi}) > \mu(E_{i+1,\xi}/E_{i,\xi})$, $\mu(E_{i,s}/E_{i-1,s} > \mu(E_{i+1,s}/E_{i,s})$ for each $s \in U$. Hence $0 \subset E_{1,s} \subset \ldots \subset E_{r,s}$ is the Harder–Narasimhan filtration of E_s. This proves the proposition.

Proposition 1.9. For any family E_S of parabolic bundles, the function $s \to \mathrm{HNP}(E_s)$ is upper semi-continuous.

Proof. Let η be the generic point of an irreducible component S_1 of S. Then by proposition 1.7 there is an open subset U of S_1 such that $\mathrm{HNP}(E_s)$ is constant for $s \in U$. Note that $\dim (S_1 - U) < \dim S_1$. If $\eta_0 \in S_1 - U$ then by Prop. 1.6, $\mathrm{HNP}(E\eta_0) \geq \mathrm{HNP}(E_\eta)$. Now the result follows by noetherian induction.

PROPOSITION 1.10. Let the base field be $\mathbf{C}$. Let E_S be a family of parabolic bundles of a constant Harder–Narasimhan type parametrized by a variety S. Then there exists a filtration of E_S which is continuous in the complex topology and which specializes to the Harder–Narasimhan filtrations at all points of S.

Proof. By proposition 1.7 there exists a Zariski open subvariety $U \subseteq S$ such that the Harder–Narasimhan filtration varies algebraically over U. Hence to prove the continuity of the filtration over all of S, it is enough to prove that it is continuous over any locally closed curve in S which intersects U. Hence we may assume that $\dim S = 1$. There is no essential difference if S is replaced by its desingularization, so we may assume that S is non-singular. Let U be an open subset of S and $0 \subset E_1 \subset \ldots \subset E_r$ a filtration of E_U which restricts to the Harder–Narasimhan filtration of E_s for any $s \in U$. By the properness over base of the appropriate Grassman bundles, the filtration $0 \subset E_1 \subset \ldots \subset E_r$ of E_U extends to a filtration of E_S over all of S which is an algebraic filtration. Since this filtration has the right Harder–Narasimhan polygon, it must restrict to the Harder–Narasimhan filtration at each $s \in S$.

Next we introduce the notion of the *compound type* (λ, I) of a parabolic bundle. The reason we introduce this is as follows. Given a fixed parabolic data and a fixed parabolic Harder–Narasimhan type λ, there can exist two parabolic bundles E and F which have the given parabolic data and the given Harder–Narasimhan type λ, but such that the parabolic data for the corresponding quotients E_i/E_{i-1} and F_i/F_{i-1} is distinct for some i. It is easy to see that the parabolic data for E_i/E_{i-1} will be identical to that for F_i/F_{i-1} all i if and only if the bundles E and F have the same *intersection matrix* I defined as follows.

DEFINITION 1.11. Let E be a parabolic bundle over X and let $E_P = F_1 \supset \ldots \supset F_m \supset 0$ be the parabolic flag over the parabolic point $P_\varepsilon X$. Let $0 \subset G_1 \subset \ldots \subset G_r = E_P$ be the filtration of the fibre E_P induced by the Harder–Narasimhan filtration of E. Then the *intersection matrix* of E is the $r \times m$ matrix I with integral entries defined by the formula

$$I_{1,m} = \dim\ (G_1 \cap F_m)$$

$$I_{p,q} = \dim\ (G_p \cap F_q) - \sum_{\substack{i \leq p \\ j \geq q}} I_{i,j} \qquad (i,j) \neq (p,q)$$

The pair (λ, I) will be called the compound type of E.

PROPOSITION 1.12. Let E_S be a family of parabolic bundles of a constant Harder–Narasimhan type. Then the subset S_I of S corresponding to any particular intersection matrix I is closed in S.

Proof. The proof follows easily from the proposition 1.8 and arguments similar to the ones used for earlier propositions.

The infinitesimal deformations of a parabolic bundle E are parametrized by the vector space $H^1(X, \text{Par End } E)$ where Par End E is the sheaf of germs of endomorphisms of E which preserve the parabolic structure. The following generalization of the lemma 15.5 of Atiyah and Bott [1], says that there exist 'sufficiently large' families of parabolic bundles.

PROPOSITION 1.13. For any parabolic bundle E_0 on X, there exists a family E_V of parabolic bundles parametrized by a non-singular variety V, and a closed point $x_0 \in V$, such that

(i) $E_{x_0} \cong E_0$ and

(ii) the infinitesimal deformation map from $T_{x_0}(V)$ to $H^1(X, \text{Par End } E_0)$ is an isomorphism.

Proof. By lemma 15.5 in Atiyah and Bott [1], there exists a family E_S of ordinary vector bundles parametrized by a non-singular variety S, and a closed point $s_0 \in S$, such that $E_{s_0} \cong E_0$, and the deformation map $T_{s_0}(S) \to H^1(X, \text{End } E_0)$ is surjective. Let $P \in X$ be the parabolic point, and let $\{E_{0,p} = F_1 \supset F_2 \supset \ldots \supset F_r \supset 0\} = F$ be the parabolic flag in the fibre $E_{0,p}$. Let $\mathcal{F}$ be the flag manifold of all flags of the above type in $E_{0,p}$. Let $E_{S\times\mathcal{F}}$ be the vector bundle on X parametrized by $S \times \mathcal{F}$ which is the pull-back of E_S. Then $E_{S\times\mathcal{F}}$ has a natural flag structure over $P \in X$ making it a family of parabolic bundles on X parametrized by $S \times \mathcal{F}$, such that $E_{(s_0, F)} \cong E_0$ as a parabolic bundle. It is easy to see that the infinitesimal deformation map

$$T_{(s_0, F)}(S \times \mathcal{F}) \to H^1(X, \text{ Par End } E_0)$$

for this family at $(s_0, F) \in S \times \mathcal{F}$ is surjective. Hence we may choose an appropriate non-singular locally closed subvariety $V \subseteq S \times \mathcal{F}$, which contains the point $x_0 = (s_0, F)$, such that the infinitesimal deformation map $T_{x_0}(V) \to H^1(X, \text{Par End } E_0)$ for the induced family E_V is an isomorphism.

Let E be a parabolic bundle, and let Par End$'$ E be the sheaf of germs of all endomorphisms of E which preserve both the parabolic filtration and the Harder–Narasimhan filtration of E. Let Par End$''$ E be the cokernel of the inclusion Par End$'$ $E \hookrightarrow$ Par End E.

PROPOSITION 1.14. For any parabolic bundle E, $H^0(X, \text{Par End}'' E) = 0$.

Proof. The successive quotients of the Harder–Narasimhan filtration are semi-stable parabolic bundles with decreasing par μ. Now if D_1 and D_2 are two semi-stable parabolic bundles with par $\mu(D_1) >$ par $\mu(D_2)$ then $\mathrm{Hom}(D_1, D_2) = 0$ by the proposition 9, Part 3 of Seshadri [8]. From this our desired result follows using induction over the length of the Harder–Narasimhan filtration.

Let E_S be a family of parabolic bundles parametrized by a variety S. For any Harder–Narasimhan type λ, let S_λ be the corresponding subset of S. Then by Prop. 1.8, S_λ is locally closed subvariety of S. Let $s_0 \in S_\lambda$ and let ϕ: $T_{s_0}(S) \to H^1(X$, Par End $E_{s_0})$ be the infinitesimal deformation map at s_0. Then the following is easy to prove and we omit the proof.

PROPOSITION 1.15. The tangent space $T_{s_0}(S_\lambda)$ to the subvariety S_λ at s_0 is the inverse image of $H^1(X$, Par End$'$ $E_{s_0}) \subseteq H^1(X$, Par End $E_{s_0})$ under the infinitesimal deformation map $\phi: T_{s_0}(S) \to H^1(X$, Par End $E_{s_0})$.

COROLLARY 1.16. If ϕ above is surjective, then the codimension of S_λ in S at s_0 equals the dimension of $H^1(X$, Par End$''$ $E_{s_0})$.

Proof. This is clear from Prop. 1.13 by which $H^0(X$, Par End$''$ $E_{s_0}) = 0$.

PROPOSITION 1.17. For a parabolic bundle E, the dimension of $H^1(X$, Par End$''$ $E)$ is given by the following formula

$$h^1(X, \text{Par End}''\, E) = \sum_{i\le j} (m_j d_i - m_i d_j) + g(X)\left((\text{rank } E)^2 - \sum_{i\le j} m_i m_j\right) - \sum_{i\le j} n_i n_j + \sum_{\substack{p\ge k\\ q\le l}} I_{pq} I_{kl},$$

where m_i and d_i are the rank and (ordinary) degree of the ith quotient of the Harder–Narasimhan filtration, and n_i is the multiplicity of the ith weight in the parabolic structure.

Proof. Let End$'$ E be the sheaf of germs of endomorphisms of E which preserve the filtration $0 \subset E_1 \subset \ldots \subset E_r$, but may or may not preserve the parabolic structure. Let End$''$ E = End E/End$'$ E. Then there is an exact sequence $0 \to$ Par End$''$ $E \to$ End$''$ $E \to \mathcal{S} \to 0$ where $\mathcal{S}$ is some skyscraper sheaf. As $h^0(X$, Par End$''$ $E) = 0$ by Prop. 1.13, $h^1(X$, Par End$''$ $E) = -\chi($Par End$''$ $E) = -\chi($End$''$ $E) + \chi(\mathcal{S})$. The Riemann–Roch theorem shows that

$$-\chi(\text{End}''\; E) = \sum_{i\le j} (m_j d_i - m_i d_j) + (g(X)-1)\left((\text{rank } E)^2 - \sum_{i\le j} m_i m_j\right),$$

while it is easy to see explicitly that $h^0(X, \mathcal{S})$ equals

$$(\text{rank } E)^2 - \sum_{i\le j} n_i n_j + \sum_{i\le j} n_i n_j + \sum_{\substack{i\le j\\ p\ge q}} I_{ip} I_{jq}$$

As $\mathcal{S}$ is a skyscraper sheaf, $h^0(X, \mathcal{S}) = \chi(\mathcal{S})$, which completes the proof.

2. Sobolev spaces

Let E be a fixed C^∞ parabolic bundle on the compact Riemann surface X, of ranl n, degree d, parabolic filtration $E_P = F_1 \supset \ldots \supset F_m \supset 0$ over the parabolic poin $P \in X$, and weights $0 \leq \alpha_1 < \ldots < \alpha_m < 1$. Let E be given a fixed Hermitia metric, and X be given a fixed Riemannian metric.

As shown by Atiyah and Bott [1], all the holomorphic structures on E form a affine space $\mathscr{C}(E)$ (or simply $\mathscr{C}$) whose tangent space is the infinite dimensiona vector space $\Omega^{0,1}$ (End E) of all C^∞ global endomorphisms of E. The points of are the operators $d'': a^0(E) \to a^{0,1}(E)$ which satisfy (i) d'' is $\mathbb{C}$-linear and (i $d''(fv) = \bar{\partial} f \otimes v + f d'' v$ where f is a local C^∞ function and v a local C^∞ section of E The holomorphic bundle corresponding to an operator d'' will be denoted by $E_{d''}$

Let Aut(E) denote the group of all C^∞-automorphisms of E, and let $\mathscr{P}$ = Pa Aut(E) be the subgroup which preserves the parabolic filtration. Then Par Aut$(E$ acts on $\mathscr{C}$, and the orbits are the isomorphism classes of holomorphic paraboli bundles on X of the given rank, degree and parabolic structure.

We want to analyze the action of $\mathscr{P}$ on $\mathscr{C}$. For this, following § 14 of Atiyah an Bott we first replace $\mathscr{C}$ by the Sobolev space $\mathscr{C}^{k-1}$ (denoted by $\mathscr{A}^{k-1}$ in [1]) of a d''-operators of class $\mathscr{H}^{k-1}$, where k is some large positive integer. Consider th Sobolev space $\mathscr{H}^k$ (End E) of all global sections of End E of class $\mathscr{H}^k$. If ϕ is a C endomorphism of E, then we can restrict ϕ to the parabolic point $P \in X$ to get continuous linear map from a^0 (End E) to End (E_p). Even though $\{P\} \subset X$ is a se of measure zero, we can extend this map to the Sobolev space $\mathscr{H}^k$ (End E) (whic contains a^0 (End E) as a subspace) provided k is large enough. This extension is th well known *trace map* $\mathscr{H}^k$ (End E) $\to$ End (E_p) (see e.g. Triebel [10] §2.7.2). As is assumed to be large enough, the trace map $\mathscr{H}^k$ (End E) $\to$ End (E_P) is define and continuous where E_P is the fibre of E over the parabolic point $P \in X$. Th kernel of this map is the Hilbert space Par $\mathscr{H}^k$ (End E) of all paraboli endomorphisms of E of class $\mathscr{H}^k$. This is then the Lie algebra of the Hilbert Li groups $\mathscr{P}^k$ of all parabolic automorphisms of E of class $\mathscr{H}^k$. Note that $\mathscr{P}^k$ i definable as a closed subgroup of the Hilbert Lie group Aut$(E)^k$ of a automorphisms of E of class k by using the trace map which is defined an continuous as k is large.

As shown by Atiyah and Bott Aut$(E)^k$ acts smoothly on $\mathscr{C}^{k-1}$ which extends th action of Aut(E) on $\mathscr{C}$. It follows that $\mathscr{P}^k$ also acts smoothly on $\mathscr{C}^{k-1}$.

For any $A \in \mathscr{C}^{k-1}$, let F: Aut$(E)^k \to \mathscr{C}^{k-1}$ be the map given by the action on i.e. $F(g) = g(A)$. Then lemma 14.6 of Atiyah and Bott says that the differential d at identity is a Fredholm operator. As the Lie algebra Par $\mathscr{H}^k$ End E of $\mathscr{P}^k$ is finite codimension in the Lie algebra $\mathscr{H}^k$ End E of Aut$(E)^k$, it follows that for th restricted map $\mathscr{P}^k \to \mathscr{C}^{k-1}$, the differential at identity is a Fredholm operato Applying the smooth group action of $\mathscr{P}^k$ on $\mathscr{C}^{k-1}$, it follows that the differential is Fredholm operator (of constant kernel and cokernel dimensions) at all points of th orbit of A under $\mathscr{P}^k$. The implicit function theorem for Banach manifolds the implies the following:

PROPOSITION 2.1. For neighbourhoods U of the identity in $\mathcal{P}^k$ and V of A in $\mathcal{C}^{k-1}$, the image $U(A)$ is a closed Banach submanifold of V of finite codimension.

Now the proof of the lemma 14.8 of Atiyah and Bott applies to prove the following:

PROPOSITION 2.2. C^∞ points are dense in every $\mathcal{P}^k$ orbit in $\mathcal{C}^{k-1}$.

The next proposition follows at once from lemma 14.9 of Atiyah and Bott [1].

PROPOSITION 2.3. Let $A, B \in \mathcal{C}$ and $g \in \mathcal{P}^k$ with $B = g(A)$. Then $g \in \mathcal{P}$, i.e. g is C^∞.

PROPOSITION 2.4. Let $A \in \mathcal{C} \subset \mathcal{C}^{k-1}$. Then the fibre of the normal bundle to the $\mathcal{P}^k$-orbit of A at A is the vector space $H^1(X, \text{Par End } E)$ where E is given the holomorphic structure corresponding to A.

Proof. An infinitesimal parabolic endomorphism $\phi \in \text{Par } \mathcal{H}^k \text{ End } E$ alters any element A of $\mathcal{C}^{k-1}$ by the addition of $d''_A\phi$, where d''_A is the operator from $\mathcal{H}^k$ End E to $\mathcal{H}^{k-1}\Omega^{0,1}$ End E corresponding to A. Here $\mathcal{H}^{k-1}\Omega^{0,1}$ End E denotes $(0,1)$-forms of class $\mathcal{H}^{k-1}$ with coefficients in End E. It is easy to see that if E is given a holomorphic structure corresponding to $A \in \mathcal{C}$, then the sequence of sheaves $0 \to \text{Par End } E \to \textit{Par } \mathcal{H}^k \textit{ End } E \xrightarrow{d''_A} \mathcal{H}^{k-1}\Omega^{0,1} \textit{ End } E \to 0$ is exact where *Par* $\mathcal{H}^k$ *End E* is the sheaf of germs of parabolic endomorphisms of E of class $\mathcal{H}^k$ and $\mathcal{H}^{k-1}\Omega^{0,1}$ *End E* has an analogous meaning. Note that the sheaves *Par* $\mathcal{H}^k$ *End E* and $\mathcal{H}^{k-1}\Omega^{0,1}$ *End E* are fine sheaves. Hence $H^1(X, \text{Par End } E)$ is the cokernel of d''_A: Par $\mathcal{H}^k$ End $E \to \mathcal{H}^{k-1}\Omega^{0,1}$ End E, which is just the normal to the $\mathcal{P}^k$-orbit at A.

The infinitesimal deformation map of an algebraic family E_S of parabolic bundles now has the following interpretation (see the proof of lemma 15.5 of Atiyah and Bott). Let $s_0 \in S$ be a smooth point, and let U be a small (euclidean) neighbourhood of s_0 in S such that the family when restricted to U is C^∞-trivial. By fixing a C^∞ isomorphism of the family E_U with the trivial C^∞ family $E_{s_0} \times U \to X \times U$, we get a C^∞ map ψ of U into $\mathcal{C} = \mathcal{C}(E_{s_0})$. The differential of ψ at s_0 is then a map of $T_{s_0}(S)$ into the tangent space to $\mathcal{C}$ at $\psi(s_0)$. Projecting onto the normal to the $\mathcal{P}$-orbit (which is $H^1(X, \text{Par End } E_{s_0})$ as seen above) then gives the infinitesimal deformation map.

The existence of 'large enough' algebraic families of parabolic bundles (see Prop. 1.13, § 1) now means the following.

PROPOSITION 2.5. For any C^∞-point $A \in \mathcal{C}^{k-1}$ the $\mathcal{P}^k$-orbit of A has as a local transversal at A a finite dimensional manifold which represents an algebraic family of parabolic bundles parametrized by a non-singular variety.

Let $\mathcal{C}_{\lambda,I}$ be the subset of $\mathcal{C}$ corresponding to a compound type (λ, I), where λ is the Harder–Narasimhan type and I the intersection matrix. Define the subset $\mathcal{C}^{k-1}_{\lambda,I}$ of $\mathcal{C}^{k-1}$ to be the set of all elements of $\mathcal{C}^{k-1}$ which are $\mathcal{P}^k$-equivalent to an

element of $\mathscr{C}_{\lambda,I}$. By Prop. 2.2 and Prop. 2.3, we see that the $\mathscr{C}^{k-1}_{\lambda,I}$ are well-defined, mutually disjoint and their union is all of $\mathscr{C}^{k-1}$.

Remark 2.6. Given any compound type (λ, I) (such that λ and I are compatible and λ is a convex polygon), there exists at least one parabolic bundle E on X of compound type (λ, I). This is so because given any rank, degree, and parabolic structure, there exists (see [4], [8], [9]) at least one semi-stable parabolic bundle of this rank, degree and parabolic structure. Hence for any compound type (λ, I) parabolic bundles corresponding to the successive quotient types of (λ, I), and then their direct sum is a parabolic bundle of compound type (λ, I). In particular, each $\mathscr{C}_{\lambda,I}$ is non-empty and hence each $\mathscr{C}^{k-1}_{\lambda,I}$ is non-empty.

Let the partial ordering on the Harder–Narasimhan types λ be extended to compound types (λ, I) by defining $(\lambda_1, I_1) < (\lambda_2, I_2)$ iff $\lambda_1 < \lambda_2$. Note that if $\lambda_1 = \lambda_2$, then unless $I_1 = I_2$, these two compound types are incomparable. Then propositions 1.9 and 1.12 imply that the compound type is upper semi-continuous in an algebraic family of parabolic bundles.

Proposition 2.7. The compound type is an upper semi-continuous function on $\mathscr{C}^{k-1}$. Further, each $\mathscr{C}^{k-1}_{\lambda,I}$ is a locally closed submanifold of $\mathscr{C}^{k-1}$ codimension given by Prop. 1.17, and $\mathscr{C}_{\lambda,I}$ is also a locally closed submanifold of $\mathscr{C}$ of the above codimension.

Proof. Let $A \in \mathscr{C} \subset \mathscr{C}^{k-1}$. Then the $\mathscr{P}^k$-orbit has a local transversal S at A which represents an algebraic family such that the infinitesimal deformation map is an isomorphism at A (though it may not be injective at surrounding points). Let $G \subset \mathscr{P}^k$ be the isotropy group at A, which is just the group of holomorphic parabolic automorphisms of E_A. Since G is finite dimensional, it has a closed complement W in a neighbourhood of identity in $\mathscr{P}^k$. Then W is a Banach manifold and the differential of the map $\alpha\colon W \times S \to \mathscr{C}^{k-1}$ (given by the group action) is an isomorphism at (e, A) where $e \in \mathscr{P}^k$ is the identity. Hence by the inverse function theorem for Banach manifolds, we may assume, after shrinking W and S suitably, that the map $\alpha\colon W \times S \to \mathscr{C}^{k-1}$ is an isomorphism of $W \times S$ with some open neighbourhood $W(S)$ of A in $\mathscr{C}^{k-1}$.

Now, the compound type is an upper semi-continuous function on S. Hence it is upper semi-continuous on $W(S)$. Further, as the infinitesimal deformation map for S is everywhere surjective it follows by 1.15, 1.16 and 1.17 that each $S_{\lambda,I}$ is a non-singular locally closed subvariety of S of codimension given by Prop. 1.17. Hence each $\mathscr{C}^{k-1}_{\lambda,I}$ is a locally closed submanifold in $W(S)$ of the correct codimension. At any C^∞ point $B \in \mathscr{C}^{k-1}_{\lambda,I}$, there exists a C^∞ complement. Hence within $W(S)$, $\mathscr{C}_{\lambda,I}$ is a locally closed submanifold of $\mathscr{C}$ of the correct codimension. As open sets of the form $W(S)$ cover $\mathscr{C}^{k-1}$, proposition 2.7 is proved.

Proposition 2.8. The semi-stable stratum $\mathscr{C}_{ss}$ is connected.

Proof. Each C^∞ point $A \in \mathscr{C}^{k-1}$ has an open neighbourhood of the form $W(S)$ as in the proof of Prop. 2.7. Let $W^\infty = W \cap \mathscr{P}$ be the set of C^∞ points of W. Then by Prop. 2.3, $W^\infty(S)$ is equal to $W(S) \cap \mathscr{C}$. By suitably shrinking W and S, we may assume that both W^∞ and S are connected. We know that the semi-stable part of S is open in the Zariski topology, hence connected and dense provided it is non-empty. Hence the semi-stable subset of $W^\infty(S)$ is connected and dense provided it is non-empty. Hence each point $A \in \mathscr{C}$ has a neighbourhood V such that $V \cap \mathscr{C}_{ss}$ is connected and dense provided it is non-empty. From this it follows at once that $\mathscr{C}_{ss}$ is connected since $\mathscr{C}$ is connected.

Let $\mathscr{F}_{\lambda, I}$ be the space of all C^∞ filtrations of E of compound type (λ, I). Fix any one such filtration, and let $\mathscr{P}_{\lambda, I} = \text{Par Aut}\,(E_{\lambda, I})$ denote the subgroup of Par Aut (E) which preserves this filtration. Then $\mathscr{F}_{\lambda, I}$ is by definition the homogeneous space $\mathscr{P}/\mathscr{P}_{\lambda, I}$. We have a canonical map $f\colon \mathscr{C}_{\lambda, I} \to \mathscr{F}_{\lambda, I}$ which associates to any holomorphic structure on E of compound type (λ, I) the associated Harder–Narasimhan filtration. Let $\mathscr{B}_{\lambda, I}$ be the fibre of f over the chosen base point in $\mathscr{F}_{\lambda, I}$. Note that $\mathscr{B}_{\lambda, I}$ is carried into itself by the action of $\mathscr{P}_{\lambda, I}$. We prove in the next section that $\mathscr{B}_{\lambda, I}$ is a manifold isomorphic to $\prod_i \mathscr{C}_{ss}(D_i) \times \prod_{i<j} \Omega^0$ Hom (D_j, D_i) where D_i are the successive quotients of the chosen filtration of E. An analogous statement holds for the space $\mathscr{B}^{k-1}_{\lambda, I}$ used in the proof of the next proposition.

PROPOSITION 2.9. The equivariant cohomology of the pair $(\mathscr{C}_{\lambda, I}, \mathscr{P})$ is isomorphic to the equivariant cohomology of $(\mathscr{B}_{\lambda, I}, \mathscr{P}_{\lambda, I})$.

Proof. The proof is the exact analogue of the proof for the corresponding statement given in Atiyah and Bott, § 14. For the sake of completeness, we sketch the steps involved. Let $\mathscr{P}^k_{\lambda, I}$ be the subgroup of $\mathscr{P}^k$ which preserves the chosen filtration of E, and let $\mathscr{F}^k_{\lambda, I} = \mathscr{P}^k/\mathscr{P}^k_{\lambda, I}$. The map $f\colon \mathscr{C}_{\lambda, I} \to \mathscr{F}_{\lambda, I}$ then extends to give a $\mathscr{P}^k$-equivariant map $f^k\colon \mathscr{C}^{k-1}_{\lambda, I} \to \mathscr{F}^k_{\lambda, I}$. The continuity of f^k follows from Prop. 1.10 which says that the Harder–Narasimhan filtration varies continuously in an algebraic family, together with the earlier results in this section about the action of $\mathscr{P}^k$ on $\mathscr{C}^{k-1}$. Now $\mathscr{P}^k_{\lambda, I}$ is a closed subgroup of the Hilbert Lie group $\mathscr{P}^k$, and hence the fibration $\mathscr{P}^k \to \mathscr{F}^k_{\lambda, I}$ is locally trivial and hence is a principal $\mathscr{P}^k_{\lambda, I}$ bundle. Then it is shown that $f^k\colon \mathscr{C}^{k-1}_{\lambda, I} \to \mathscr{F}^k_{\lambda, I}$ is $\mathscr{P}^k$-equivariantly isomorphic to the associated bundle to $\mathscr{P}^k \to \mathscr{F}^k_{\lambda, I}$ with fibre $\mathscr{B}^{k-1}_{\lambda, I}$, where $\mathscr{B}^{k-1}_{\lambda, I}$ is the fibre of f^k over the chosen base point in $\mathscr{F}_{\lambda, I}$. Hence from § 13 of Atiyah and Bott it follows that the equivariant cohomologies of the pairs $(\mathscr{C}^{k-1}_{\lambda, I}, \mathscr{P}^k)$ and $(\mathscr{B}^{k-1}_{\lambda, I}, \mathscr{P}^k_{\lambda, I})$ are isomorphic. Finally, an appeal to the 'standard approximation theorems' which say that the homotopy properties of the various function spaces are all independent of k (see Palais [6], theorem 13.14) shows that for equivariant cohomology the pairs $(\mathscr{C}_{\lambda, I}, \mathscr{P})$ and $(\mathscr{C}^{k-1}_{\lambda, I}, \mathscr{P}^k)$ are equivalent and $(\mathscr{B}_{\lambda, I}, \mathscr{P}_{\lambda, I})$ and $(\mathscr{B}^{k-1}_{\lambda, I}, \mathscr{P}^{k-1}_{\lambda, I})$ are equivalent. Hence the pairs $(\mathscr{C}_{\lambda, I}, \mathscr{P})$ and $(\mathscr{B}_{\lambda, I}, \mathscr{P}_{\lambda, I})$ are equivalent for equivariant cohomology.

3. Betti numbers of the moduli space

3.1. Atiyah and Bott prove the following result which is basic for our computation (see §1 of [1]). Let M be a (possibly infinite dimensional) connected manifold together with a given action of a group G. Let $M = \bigcup_{j \in J} M_j$ be a G-invariant stratification of M by countably many locally closed submanifolds M_j which are connected and have finite codimensions c_j. For each integer q let there be only finitely many M_j such that $c_j < q$. Let the indexing set J of the stratification have a partial ordering such that (i) the indexing function $M \to J$ is upper semi-continuous (ii) J has a minimum $j_0 \in J$ (iii) for every finite subset $I \subseteq J$ there exist a finite number of minimal elements of the complement of I such that any other element of the complement is greater than at least one of them. Let the normal N_j of any stratum M_j be orientable. Let the equivariant Euler class of N_j not be a zero-divisor in the equivariant cohomology $H^*_G(M_j, \mathbf{Z})$. Then the stratification is equivariantly perfect over $\mathbf{Z}$. In particular the various equivariant Poincaré series are related by

$$\bar{P}_t(M) = \Sigma t^{c_j} \bar{P}_t(M_j).$$

We now proceed to apply the above to our stratification $\mathscr{C} = \cup \mathscr{C}_{\lambda, I}$ to deduce the equivariant cohomology of the semi-stable stratum. We begin with the equivariant cohomology of ($\mathscr{C}$, Par Aut (E)).

PROPOSITION 3.2. Let $\mathscr{F}$ be the flag variety of flags in $\mathbb{C}^n$ of the type specified by the parabolic structure, and let B Aut (E) be the classifying space of the group Aut (E). Then the Par Aut (E)-equivariant cohomology of $\mathscr{C}$ is given by

$$H^*{}_{\mathrm{Par\ Aut}\,(E)}(\mathscr{C}) \approx H^*(B\ \mathrm{Aut}\,(E)) \otimes H^*(\mathscr{F}).$$

Proof. Since $\mathscr{F} \approx \mathrm{Aut}\,(E)/\mathrm{Par\ Aut}\,(E)$, we have a fibration

$$\mathscr{F} \to B\ \mathrm{Par\ Aut}\,(E) \to B\ \mathrm{Aut}\,(E).$$

Let P be the parabolic subgroup of $GL(n, \mathbb{C})$, where $n = \mathrm{rank}\ E$, which preserves a flag of the given type. Then the above fibration is a pull-back of the fibration

$$\mathscr{F} \to BP \to BGL(n, \mathbb{C}).$$

Now using theorem 20.6 of Borel [2] it follows easily that the latter fibration is cohomologically trivial (with $\mathbf{Z}$-coefficients). Hence the first fibration is also cohomologically trivial. Hence we get

$$H^*(B\ \mathrm{Par\ Aut}\,(E)) \approx H^*(B\ \mathrm{Aut}\ E) \otimes H^*(\mathscr{F}).$$

As $\mathscr{C}$ is contractible, the equivariant cohomology of $\mathscr{C}$ is the cohomology of B Par Aut (E). Hence

$$H^*{}_{\mathrm{Par\ Aut}\,(E)}(\mathscr{C}) \approx H^*(B\ \mathrm{Aut}\ E) \otimes H^*(\mathscr{F}).$$

Remark 3.3. Atiyah and Bott have determined the cohomology of B Aut (E) (see theorem 2.15 in [1]). In particular, they prove that B Aut (E) is torsion-free. Hence it follows from Prop. 3.1 that $H^*_{\text{Par Aut }(E)}(\mathscr{C})$ has no torsion.

We next reduce the equivariant cohomology of any non semi-stable stratum $\mathscr{C}_{\lambda, I}$ to the tensor product of the equivariant cohomologies of semi-stable strata for some parabolic bundles of lower ranks. This is the basic inductive step.

PROPOSITION 3.4. Let $0 \subset E_1 \subset E_2 \subset \ldots \subset E_r = E$ be a fixed C^∞-filtration of E of type (λ, I). Then

$$H^*_{\text{Par Aut }(E)}(\mathscr{C}_{\lambda, I}) \approx \bigotimes_{i=1}^{r} H^*_{\text{Par Aut }(D_i)}(\mathscr{C}_{ss}(D_i)),$$

where $D_i = E_i/E_{i-1}$ with the induced parabolic structure.

Proof. As shown in §2, the pairs $(\mathscr{C}_{\lambda, I}$, Par Aut $(E))$ and $(\mathscr{B}_{\lambda, I}$, Par Aut $(E_{\lambda, I}))$ are equivalent for equivariant cohomology. Now fix a C^∞ parabolic splitting $E \approx D_1 \oplus \ldots \oplus D_r$ of the chosen filtration $0 \subset E_1 \subset E_2 \subset \ldots \subset E_r$ of type (λ, I), so that $E_i = D_1 \oplus \ldots \oplus D_i$. Let Par Aut $(E^0_{\lambda, I})$ and $\mathscr{B}^0_{\lambda, I}$ be the automorphisms and complex structures in $\mathscr{B}_{\lambda, I}$ compatible with this decomposition. Then we have

$$\text{Par Aut }(E^0_{\lambda, I}) \approx \Pi \text{ Par Aut }(D_i)$$

$$\mathscr{B}^0_{\lambda, I} \approx \Pi \mathscr{C}_{ss}(D_i).$$

(Note that the parabolic structure on D_i here is by definition the parabolic structure of E_i/E_{i-1}). On the other hand the natural homomorphism

$$\text{Par Aut }(E_{\lambda, I}) \rightarrow \text{Par Aut }(E^0_{\lambda, I})$$

is a homotopy equivalence, and the fibration $\mathscr{B}_{\lambda, I} \rightarrow \mathscr{B}^0_{\lambda, I}$ is the globally trivial fibration with fibre the infinite dimensional vector space $\bigoplus_{i<j}$ Hom (D_j, D_i) and is compatible with the group action. It follows that, for purposes of equivariant cohomology, the pairs $(\mathscr{C}_{\lambda, I}$, Par Aut $(E))$ and $(\mathscr{B}_{\lambda, I}$, Par Aut $(E^0_{\lambda, I}))$ are equivalent. Hence we see that the equivariant cohomology of the stratum $\mathscr{C}_{\lambda, I}(E)$ is isomorphic to the tensor product of the equivariant cohomology of the semi-stable strata for the quotients D_i. This is the exact analogue of Prop. 7.12 of Atiyah and Bott.

PROPOSITION 3.5. Each stratum $\mathscr{C}_{\lambda, I}$ is connected.

Proof. We have already proved that the semi-stable stratum is connected (see §2). Hence $\mathscr{B}^0_{\lambda, I} \approx \Pi \mathscr{C}_{ss}(D_i)$ is connected. Hence $\mathscr{B}_{\lambda, I}$ is connected as it is a fibration over $\mathscr{B}^0_{\lambda, I}$ with a vector space as the fibre. Now we have a fibration $\mathscr{C}_{\lambda, I} \rightarrow \mathscr{F}_{\lambda, I}$ with fibre $\mathscr{B}_{\lambda, I}$. Hence it remains to show that $\mathscr{F}_{\lambda, I}$ is connected. For this we need the following which is easy to prove.

LEMMA. Let G be a topological group, let G', H be closed subgroups, and let $H' = G' \cap H$. Let G/G' be simply connected and let G/H and H/H' be connected. Then G'/H' is connected.

Now take $G' = \mathrm{Aut}(E)$, $G' = \mathrm{Par\ Aut}(E)$, and H to be the subgroup of G which preserves the filtration $0 \subset E_1 \subset E_2 \subset \ldots \subset E_r$ but not necessarily the parabolic structure. Atiyah and Bott show in § 9 of [1] that G/H is connected. Now, $H' = \mathrm{Par\ Aut}(E_{\lambda,I})$, and both G/G' and H/H' are homomorphic to the flag variety $GL(n)/P$ which is simply connected. It follows that $\mathscr{F}_{\lambda,I} = G'/H'$ is connected.

PROPOSITION 3.6. The equivariant Euler class of the normal bundle $N_{\lambda,I}$ to a stratum $\mathscr{C}_{\lambda,I}$ is not a zero divisor in the equivariant cohomology ring $H^*_{\mathrm{Par\ Aut}(E)}(\mathscr{C}_{\lambda,I})$ with integral coefficients.

Proof. The equivariant Euler class of $N_{\lambda,I}$ is by definition the usual Euler class of the vector bundle $EG \underset{G}{\times} N_{\lambda,I}$ on the space $EG \underset{G}{\times} \mathscr{C}_{\lambda,I}$ where G denotes the group Par Aut(E). Exactly the same reduction as in the proof of Prop. 3.4 shows that we can replace the triple. (Par Aut(E), $\mathscr{C}_{\lambda,I}$, $N_{\lambda,I}$) by the triple (Par Aut$(E^0_{\lambda,I})$, $\mathscr{B}^0_{\lambda,I}$, $N^0_{\lambda,I}$) where $N^0_{\lambda,I}$ is the restriction of $N_{\lambda,I}$ to $\mathscr{B}^0_{\lambda,I}$. Now as Par Aut$(E^0_{\lambda,I}) \approx \Pi$ Par Aut(D_i), it contains the r-dimensional torus T^r, which acts trivially on $\mathscr{B}^0_{\lambda,I}$. Now at a point of $\mathscr{B}^0_{\lambda,I}$ our bundle E is a holomorphic direct sum of the D_i, and hence Par End$''E \approx \bigoplus_{i<j}$ Par Hom (D_i, D_j). Now on Par Hom (D_i, D_j) the element $(t_1, \ldots, t_r) \in T^r$ acts by $t_i^{-1}t_j$ and so it acts by the same character on H^1 $(X$, Par Hom $(D_i, D_j))$. Since the fibre of the normal bundle N is H^1 $(X$, Par End$''$ $E)$, it follows that the representation of T^r on a fibre of N is 'primitive', i.e., its one-dimensional components in a direct sum decomposition are indivisible elements of the character group of T^r. Hence by Prop. 13.4 of Atiyah and Bott it follows that the equivariant Euler class of $N_{\lambda,I}$ is not a zero-divisor in the equivariant cohomology of $\mathscr{C}_{\lambda,I}$.

With this we see that all the conditions stated in §3.1 are satisfied by the stratification $\mathscr{C} = \cup \mathscr{C}_{\lambda,I}$. Hence we have the following formula relating the various equivariant Poincaré series

(3.7) $$\bar{P}_t(\mathscr{C}) = \Sigma t^{C_{\lambda,I}} \bar{P}_t(\mathscr{C}_{\lambda,I}).$$

Substituting the result for $\bar{P}_t(\mathscr{C})$ from Prop. 3.2, we get the following formula in which the summation on the right hand side is taken only over all non semi-stable strata.

PROPOSITION 3.8. The equivariant Poincaré series for the semi-stable stratum is as given by the following inductive formula

$$\bar{P}_t(\mathscr{C}_{ss}) = P_t(\mathscr{F}) P_t(B\ \mathrm{Aut}\, E) - \Sigma t^{2d_{\lambda,I}} \bar{P}_t(\mathscr{C}_{\lambda,I})$$

where $P_t(\mathscr{F})$ is the Poincaré series for the flag variety $\mathscr{F}$ (which is well known), $P_t(B\ \mathrm{Aut}\, E)$ is the Poincaré series for B Aut E (which is known by theorem 2.15 of Atiyah and Bott [1]), $d_{\lambda,I}$ is the complex codimension of $\mathscr{C}_{\lambda,I}$ in $\mathscr{C}$ which is as given

by Prop. 1.17, and $\bar{P}_t(\mathscr{C}_{\lambda, I})$ is the equivariant Poincaré series for the stratum $\mathscr{C}_{\lambda, I}$ which is in turn expressible in terms of the Poincaré series for the semi-stable strata for parabolic bundles of lower ranks using Prop. 3.4.

PROPOSITION 3.9. Let S be the moduli space of *stable* parabolic bundles of the given rank, degree and parabolic data. Let $f\colon \mathscr{C}_s \to S$ be the canonical set-theoretic map from the stable part of $\mathscr{C}$ to S. Then f induces a homeomorphism

$$\bar{f}\colon \mathscr{C}_s/\text{Par Aut}\,(E) \xrightarrow{\simeq} S.$$

Proof. The universal property of S implies that the restriction of f to any algebraic family contained in $\mathscr{C}_s$ is continuous. From the discussion in §2 it now follows that f is itself continuous. Since the fibres of f are precisely the orbits of Par Aut (E) in $\mathscr{C}_s$, we get a continuous bijection $\bar{f}\colon \mathscr{C}_s/\text{Par Aut}\,(E) \to S$. Now, S is a non-singular quasi-projective variety, and for any local algebraic family V in $\mathscr{C}_s$ which is transversal to the orbits, the map $f\colon V \to S$ is of maximal rank and hence open. From this it follows that f is open and hence $\bar{f}$ is a homeomorphism.

THEOREM 3.10. The Poincaré series for the moduli space S of parabolic bundles over a curve in the case 'stable = semi-stable' is as follows.

$$P_t(S) = (1-t^2)\cdot \bar{P}_t(\mathscr{C}_{ss}),$$

where $\bar{P}_t(\mathscr{C}_{ss})$ is given by Prop. 3.8. Further, if the rank and degree of the parabolic bundles are coprime, then the integral cohomology of S is torsion free.

Proof. Let $\bar{\mathscr{P}}$ be the quotient group of $\mathscr{P} = \text{Par Aut}\,(E)$ by the subgroup $\mathbb{C}^*$. Let $\overline{\text{Aut}\,(E)}$ be the quotient of Aut (E) by $\mathbb{C}^*$. Atiyah and Bott show that the fibration

$$B\mathbb{C}^* \to B\,\text{Aut}\,(E) \to B\overline{\text{Aut}\,(E)}$$

is trivial in rational cohomology (and if in addition, (rank, degree) = 1, then it is trivial in integral cohomology). Hence by base change, the fibration

$$B\mathbb{C}^* \to B\,\text{Par Aut}\,(E) \to B\bar{\mathscr{P}}$$

is trivial in rational cohomology and also trivial in integral cohomology when (rank, degree) = 1. Hence it follows that for rational coefficients,

$$H^*_{\text{Par Aut}\,(E)}(\mathscr{C}_{ss}) \approx H^*_{\bar{\mathscr{P}}}(\mathscr{C}_{ss}) \otimes H^*(B\mathbb{C}^*)$$

which also holds for integral coefficients whenever (rank, degree) = 1. Now, every stable bundle is simple (see Seshadri [8]). Hence $\bar{\mathscr{P}}$ acts freely on $\mathscr{C}_{ss}$ and hence $H^*_{\bar{\mathscr{P}}}(\mathscr{C}_{ss})$ is just the singular cohomology of the moduli space. Since $P_t(B\mathbb{C}^*) = 1/(1-t^2)$ we get

$$P_t(S) = (1-t^2)\ \bar{P}_t(\mathscr{C}_{ss}).$$

Note that Prop. 3.8 gives $\bar{P}_t(\mathscr{C}_{ss})$ inductively in terms of the $\bar{P}_t(\mathscr{C}_{ss})$ for lower

ranks. To begin the process we need to know $H^*_{\text{Par Aut }(L)}(\mathscr{C}_{ss}(L))$ for a parabolic line bundle L. But in this case, the moduli space is just the Jacobian J and the rank and degree of L are coprime. Hence

$$H^*_{\text{Par Aut }(L)}(\mathscr{C}_{ss}(L)) \approx H^*(J) \otimes H(B\mathbb{C}^*)$$

with integral coefficients. Hence $H^*_{\text{Par Aut }(L)}(\mathscr{C}_{ss}(L))$ is torsion free and

$$\bar{P}_t(\mathscr{C}_{ss}(L)) = \frac{(1+t)^{2g}}{(1-t^2)}.$$

Since the stratification $\mathscr{C} = \cup \mathscr{C}_{\lambda, I}$ is equivariantly perfect over $\mathbf{Z}$, it now follows inductively that $H^*_{\text{Par Aut }(E)}(\mathscr{C}_{ss}(E))$ is torsion free for any E. Hence when (rank, deg) = 1, it follows from

$$H^*_{\text{Par Aut }(E)}(\mathscr{C}_{ss}, \mathbf{Z}) \approx H^*(S, \mathbf{Z}) \otimes H^*(B\mathbb{C}, \mathbf{Z})$$

that $H^*(S, \mathbf{Z})$ is also torsion-free.

Remark 3.11. Following exactly a similar argument used by Atiyah and Bott, one may show that the Poincaré series $P_t(S_0)$ of the moduli $S_0 \subset S$ of parabolic bundles of a fixed determinant in the 'stable = semi-stable' case is given by the formula

$$P_t(S_0) = \frac{P_t(S)}{(1+t)^{2g}}.$$

Finally, as an example we determine the Poincaré polynomial of the moduli space when rank = 2. Let $P \in X$ be the parabolic point and $0 \leq \alpha_1 < \alpha_2 < 1$ be the weights for the parabolic flag in E_p of the type $E_P = F_1 \supset F_2 \supset 0$ where dim $F_2 = 1$. Let d be the fixed degree. It is easy to see that any semi-stable bundle of this type is automatically stable.

Proposition 3.11. In the above case, the Poincaré polynomial of the moduli space is

$$P_t(S) = \frac{(1+t)^{2g-2}((1+t^3)^{2g} - t^{2g}(1+t)^{2g})}{(1-t)^2}.$$

(Note: It is easy to see directly that this rational function is indeed a polynomial).

Proof. Let E be a non-semi-stable bundle of the above type and let $0 \subset E_1 \subset E$ be its parabolic Harder–Narasimhan filtration. Let

$$k = \deg F_1 \text{ and } e = \dim((E_1)_p \cap F_2).$$

The pair (k, e) characterizes the compound type (λ, I) of E as can be seen immediately from the definition of compound type. Also note that Par $\mu(E_1) >$ Par $\mu(E)$ if and only if $2k - d + e \geq 1$. Hence we may write the

stratification of $\mathscr{C}$ as

$$\mathscr{C} = \mathscr{C}_{ss} \cup \left(\bigcup_{(k,e)} \mathscr{C}_{k,e}\right)$$

where $2k-d+e \geq 1$, $k \in \mathbf{Z}$, $e = 0, 1$. From Prop. 1.17 it follows that the complex codimension of the stratum $\mathscr{C}_{k,e}$ is $d_{k,e} = 2k-d+e-1+g$ where $g =$ genus (X). Also, it follows from Prop. 3.4 and our knowledge of $\bar{P}_t(\mathscr{C}_{ss}(L))$ for a line bundle L that

$$\bar{P}_t(\mathscr{C}_{k,e}) = \bar{P}_t(\mathscr{C}_{ss}(L))^2 = \frac{(1+t)^{4g}}{(1-t^2)^2}.$$

In our case, the flag variety $\mathscr{F}$ is just $\mathbf{P}^1$, so $P_t(\mathscr{F}) = 1+t^2$. Moreover, by theorem 2.15 of [1],

$$P_t(B \text{ Aut } E) = \frac{(1+t)^{2g}(1+t^3)^{2g}}{(1-t^2)^2(1-t^4)}.$$

Hence Prop. 3.8 gives, after some simplification,

$$\bar{P}_t(\mathscr{C}_{ss}) = \frac{(1+t)^{2g-2}((1+t^3)^{2g} - t^{2g}(1+t)^{2g})}{(1-t^2)(1-t)^2}$$

An application of theorem 3.10 now completes the proof.

Acknowledgements

The author thanks Prof. C. S. Seshadri for his constant encouragement and V. B. Mehta, T. R. Ramadas and especially A. Ramanathan for their valuable help.

References

[1] Atiyah M F and Bott R, The Yang-Mills equations on a Riemann surface, *Philos. Trans. R. Soc. London* **A308** (1982) 523
[2] Borel A, Topics in the homology theory of fibre bundles, *Lecture notes in mathematics* 36, Springer-Verlag, Berlin, Heidelberg, New York (1967)
[3] Langton S G, Valuative criteria for families of vector bundles on algebraic varieties, *Ann. Math.* **101** (1975) 88
[4] Mehta V B and Seshadri C S, Moduli of vector bundles on curves with parabolic structures, *Math. Ann.* **248** (1980) 205
[5] Narasimhan M S and Ramanathan A, Openness of the stability condition on vector bundles (1976), Unpublished manuscript, TIFR
[6] Palais R, *Foundations of global nonlinear analysis,* Benjamin, New York (1965)
[7] Seshadri C S, Moduli of vector bundles with parabolic structures, *Bull. Am. Math. Soc.* **83** (1977)
[8] Seshadri C S, Fibres Vectoriels sur les courbes Algebriques, *Asterisque* **96** (1982)
[9] Shatz S S, The decomposition and specialization of algebraic families of vector bundles, *Comp. Math.* **35** (1977) Fasc. 2, 163
[10] Triebel H, *Theory of function spaces,* Birkhauser Verlag, Basel, Boston, Stuttgart (1983)

Proc. Indian Acad. Sci. (Math. Sci.), Vol. 95, No. 2, December 1986, pp. 79–96.

Stochastic evolution equations in locally convex space

S L YADAVA
T.I.F.R. Centre, Post Box 1234, Bangalore 560 012, India

MS received 14 May 1984; revised 20 September 1985

Abstract. Ito's stochastic integral is defined with respect to a Wiener process taking values in a locally convex space and Ito's formula is proved. Existence and uniqueness theorem is proved in a locally convex space for a class of stochastic evolution equations with white noise as a stochastic forcing term. The stochastic forcing term is modelled by a locally convex space valued stochastic integral.

Keywords. Locally convex space; Wiener process; stochastic integral; Ito's formula; stochastic evolution equation.

List of symbols

Let E and F be two sequentially complete locally convex space and $(\Omega, \mathcal{B}, P)$ be a complete finite measure space. $T \in \mathbf{R}$, $T < \infty$ and $r \geq 1$.

$\mathcal{P}_E \equiv$ family of all continuous semi-norms on E.

$L(E, F) \equiv$ space of all continuous linear operators from E into F.

$L(E, E)$ will be denoted by $L(E)$.

$E^* \equiv$ space of all continuous linear functionals on E.

$L^r(\Omega, E, P) \equiv$ space of all Bochner integrable functions X from Ω into E such that $\int_\Omega \{p(X)\}^r \, \mathrm{d}P < \infty$ for every $p \in \mathcal{P}_E$.

$C([0,T], E) \equiv$ space of all continuous functions from $[0,T]$ into E.

$\mu \equiv$ Lebesgue measure on $[0,T]$.

$\mu(\mathrm{d}t)$ will be denoted by $\mathrm{d}t$.

Let A be a densely defined linear operator from E into E.

$\rho(A) \equiv$ resolvent set of A.

$R(\lambda, A) \equiv (I - \lambda A)^{-1}$ (inverse of the operator $(I - \lambda A)$) for $\frac{1}{\lambda} \in \rho(A)$.

1. Introduction

Stochastic analysis in infinite dimensions appears in several fields e.g. random vibrations, random particle system, etc. An abstract theory of stochastic evolution equation has been established by Curtain [2], Curtain and Pritchard [3], Ichikawa [5] and others in Hilbert space, Kuo [8] in abstract Wiener space. However it is also

interesting from the viewpoint of applications to discuss stochastic evolution equation in some locally convex space. In this paper, we consider the following class of stochastic evolution equation

$$du(t) = A(t)\, u(t)\, dt + \phi(t) d\omega(t) + f(t)\, dt, \quad 0 < t \leq T \tag{1}$$

$$u(0) = u_0$$

in a locally convex space F, where $A(t)$ for each $t \in [0, T]$ is a linear unbounded operator on F, $\{\omega(t), 0 \leq t \leq T\}$ a Wiener process in a locally convex space E, $\phi(t)$ and $f(t)$ are stochastic processes with values in $L(E,F)$ and F respectively. In §2 of this paper, we will study Bochner integrals, evolution operators and Wiener process in a locally convex space. In §3, we will define stochastic integrals in F with respect to a Wiener process in E and will describe some of its properties. We will also prove Ito's formula in §3. In the last section, we will prove the existence and uniqueness theorem for the stochastic evolution equation (1).

2. Notations and preliminary results

2.1 *Bochner integration in locally convex space*

Let $(\Omega, \mathcal{B}, P)$ be a complete finite measure space and E be a real sequentially complete locally convex space. Let $\mathcal{P}_E$ be the collection of all continuous semi-norms on E. A function X from Ω into E is said to be Bochner measurable if there exists a sequence $(X_n)_{n \in N}$ of simple functions from Ω to E such that X_n converges to X in E. If X is Bochner measurable then $p(X)$ is measurable for all $p \in \mathcal{P}_E$. Let

$$X = \sum_{i=1}^{n} \alpha_i | E_i$$

be a simple function. For $B \in \mathcal{B}$, define

$$\int_B X\, dP = \sum_{i=1}^{n} \alpha_i\, P(B \cap E_i).$$

A function X from Ω into E is said to be Bochner integrable if there exists a sequence $(X_n)_{n \in N}$ of simple functions converging to X in E a.s. such that

$$\lim_{n \to \infty} \int_\Omega p(X_n - X)\, dP = 0$$

for every $p \in \mathcal{P}_E$. For a Bochner integrable function X and $B \in \mathcal{B}$, define

$$\int_B X\, dP = s - \lim_{n \to \infty} \int_B X_n\, dP.$$

For a Bochner integrable function X, we have

(a) $$p\left(\int_B X \, dP\right) \leq \int_B p(X) \, dP$$

for every $p \in \mathcal{P}_E$.

(b) Let A be a closed linear operator with domain $\mathcal{D}(A) \subset E$ (E is a Fréchet space) and range in a Fréchet space Y. Let X from Ω into E be a Bochner integrable function such that $X(\Omega) \subset \mathcal{D}(A)$ and AX is also Bochner integrable, then

$$\int_B X \, dP \in \mathcal{D}(A) \tag{2}$$

and

$$A \int_B X \, dP = \int_B AX dP \tag{3}$$

for every $B \in \mathcal{B}$. We will say that the two Bochner integrable function are equal if they are equal a.s. For $1 \leq r < \infty$, define

$$L^r(\Omega, E, P) = \{X:\Omega \to E, X \text{ Bochner integrable and}$$

$$\int_\Omega \{p(X)^r \, dp < \infty \ \forall \, p \in \mathcal{P}_E\}.$$

$L^r(\Omega, E, P)$ becomes sequentially complete locally convex space under the family of semi-norms

$$q(X) = q(X, p) = [\int_\Omega \{p(X)\}^r \, dP]^{1/r},$$

where $p \in \mathcal{P}_E$. Let $(\Omega, \mathcal{B}, P)$ be a complete probability space. A Bochner integrable function X from Ω into E is called a random variable.

For a random variable X, $E(X)$ will denote the $\int X \, dP$ and will be called expectation of X. For every x^* belongs to E^* (space of all continuous linear functionals on E) we have

$$E\langle X, x^*\rangle = \langle E(X), x^*\rangle.$$

Let $\mathcal{B}'$ be a sub σ-algebra of $\mathcal{B}$ and X be an E-valued simple random variable defined on $(\Omega, \mathcal{B}, P)$ such that $X = \sum_{i=1}^{n} \alpha_i | E_i$. Define

$$E(X/\mathcal{B}') = \sum_{i=1}^{n} \alpha_i \, P(E_i/\mathcal{B}). \tag{4}$$

Obviously $E(X/\mathcal{B}')$ is a $\mathcal{B}'$-measurable E-valued random variable. $E(X/\mathcal{B}')$ will be called the conditional expectation of X given $\mathcal{B}'$. The following properties of $E(X/\mathcal{B}')$ can be easily obtained.

(a) for every $p \in \mathcal{P}_E$,

$$p(E(X/\mathcal{B}')) \leq E(p(X)/\mathcal{B}'),$$

(b) $E(E(X/\mathcal{B}')) = E(X)$,

(c) for every $x^* \in E^*$,

$$\langle E(X/\mathcal{B}'), x^*\rangle = E(\langle X, x^*\rangle/\mathcal{B}').$$

Let X belong to $L^1(\Omega, E, P)$ then there exists a sequence $\{X_n\}$ of E-valued simple random variables such that X_n converges to X in E a.s. and

$$\lim_{n\to\infty} \int_\Omega p(X_n - X)\, \mathrm{d}p = 0$$

for every $p \in \mathcal{P}_E$. Define

$$E(X/\mathcal{B}') = \lim_{n\to\infty} E(X_n/\mathcal{B}') \tag{5}$$

in $L^1(\Omega, E, P)$. Since $L^1(\Omega, E, P)$ is complete, we have $E(X/\mathcal{B}')$ is an E-valued $\mathcal{B}'$-measurable, unique upto equivalence class, random variable. $E(X/\mathcal{B}')$ will be called conditional expectation of X given $\mathcal{B}'$. Properties (a), (b), (c) of conditional expectation stated earlier for simple random variable are also obviously true for any X in $L^1(\Omega, E, P)$. For analogous result in Banach spaces see [9].

2.2 *Evolution operator in locally convex space*

Definition (2.1).

Let F be a locally convex space and $T = [0, T]$ a real finite interval and denote $\Delta(T) = \{(t, s),\ 0 \leq s \leq t \leq T\}$. A function $U(.,.)$ from $\Delta(T)$ into $L(F)$ (space of all continuous linear operator from F into F) is said to be almost strong evolution operator if

(a) $U(t, r)\ U(r, s) = U(t, s)$ for $0 \leq s \leq r \leq t \leq T$,

(b) $U(t, s)$ is strongly continuous in s on $[0, t]$ and in t on $[s, T]$.

(c) for each $t \in T$ there exists a densely-defined closed linear operator $A(t)$ on F such that

$$U(t, s) : \mathcal{D}(A(s)) \to \mathcal{D}(A(t)) \text{ for } t > s$$

(d) $\displaystyle\int_s^t A(r)\ U(r, s)x\ \mathrm{d}s = (U(t, s) - I)x$

for $x \in \mathcal{D}_{s,t}(A) = \{x \in F : U(r, s)x \in \mathcal{D}(A(r))$ for $s \leq r \leq t\}$.
Condition (d) implies that

(d′) $\dfrac{\partial}{\partial t} U(t, s)x = A(t)\, U(t, s)x$ a.s. (Lebesgue measure in $[0, T]$) for every $x \in \mathcal{D}_{s,t}(A)$.

$\{A(t), t \varepsilon [0, T]$ is called generator associated to almost strong evolution operator $U(.,.)$.

Regarding the existence of almost strong evolution operator, we have the following theorem from Yosida [12].

Theorem 2.1.
Let F be a sequentially complete locally convex space and $A(t)$ for each t, be a closed linear operator with domain $\mathcal{D}(A(t))$ and range $R(A(t))$ both in F. Suppose that $\{A(t), t \in [0, T]\}$ satisfies the following conditions:

(1) $\mathcal{D}(A(t))$ is independent of t and it is dense in F.

(2) $(0, \infty) \subset \rho(A(t))$ (resolvent set of $A(t)$) for each t. There exists a fundamental family $\mathcal{P}_F$ of continuous semi-norms on F satisfying the following condition:

for every $\lambda > 0$, $p \in \mathcal{P}_F$, there exists a positive constant M such that

$$p\{R(\lambda, A(t_n))\, R(\lambda, A(t_{n-1}))\ldots R(\lambda, A(t_1))x\} \leq M\, p(x).$$

Here M is independent of λ, $t_i (1 \leq i \leq n)$, n and x; $R(\lambda, A(t)) = (I - \lambda\, A(t))^{-1}$ and $0 \leq t_1 \leq \ldots \leq t_n = T$ is a partition of $[0, T]$.

(3) $A(s)^{-1} \in L(F)$, $A(t)\, A^{-1}(s) \in L(F)$, $0 \leq s$, $t \leq T$,

(4) For every $x \in F$,

$$\frac{1}{(t-s)}\, c(t, s)x = \frac{1}{(t-s)}\, (A(t)\, A^{-1}(s) - I)x$$

is bounded and uniformly continuous in t and s, $t \neq s$ and

$$\lim_{k \to \infty} k\, c\left(t, t - \frac{1}{k}\right)x = c(t)x$$

exists uniformly in t, where $c(t) \in L(F)$. We have moreover,

$$p\left(\frac{1}{(t-s)}\, c(t, s)x\right) \leq Np(x)$$

for some constant $N > 0$ independent of $x \in F$, t and s. Then $\{A(t), 0 \leq t \leq T\}$ is generator of an almost strong evolution operator.

Let μ denote Lebesgue measure on $[0, T]$ and as usual, we will denote $\mu(\mathrm{d}t)$ by d
For the existence and uniqueness of the solution of deterministic non-homogeneou
evolution equation, we have

Proposition 2.2

Let F be a sequentially complete locally convex space and $\{A(t), 0 \leq t \leq T\}$ be th
generator of almost strong evolution operator $U(.,.)$. Assume that

(1) $U(t,.)f(.) \in \mathfrak{D}(A(t))$ for almost all t,

(2) $f \in L^1([0, T], F, \mu)$ and for each t, $A(t)\, U(t,.)f(.) \in L^1([0, T], F, \mu)$

Then the abstract evolution equation

$$\dot{u}(t) = A(t)\, u(t) + f(t),$$

$$u(s) = u_0 \in \mathfrak{D}(A(s)), \tag{6}$$

in F has a unique strongly continuous solution

$$u(t) = U(t, s)u_0 + \int_s^t U(t, r) f(r)\, \mathrm{d}r. \tag{7}$$

2.3 *Wiener process in locally convex spaces*

Let $(\Omega, \mathfrak{B}, P)$ be a complete probability space and E be a reflexive real Fréch
space with a Schauder basis $\{e_n\}_{n \in N}$ such that for every $p \in \mathcal{P}_E$ there exists
constant M_p satisfying $p(e_n) \leq M_p$ for all n. Thus for every $x \in E$ we have

$$x = \sum_{n=1}^{\infty} \langle x, e_n^* \rangle e_n,$$

where $e_n^* \in E^*$ (strong dual of E). Further assume that $\{e_n^*\}$ be a Schauder basis
E^*. For $x_1, x_2 \in E$, define a linear operator $x_1 o\, x_2$ from E^* into E by

$$x_1 o\, x_2(x^*) = \langle x_2, x^* \rangle x_1$$

for all $x^* \in E^*$. Obviously $x_1 o\, x_2$ defines a continuous linear operator from E^* in
E.

Definition 2.2.

Let u be an E-valued random variable defined on $(\Omega, \mathfrak{B}, P)$ and $u \in L^2(\Omega, E,$
The covariance of u denoted by $\mathrm{cov}(u)$ is a linear operator from E^* into E, defin
by

$$\mathrm{cov}(u) = E\{(u - E(u))\, o\, (u - E(u))\}.$$

Since $u \in L^2(\Omega, E, P)$, $\mathrm{cov}(u)$ exists as a bounded linear operator from E^* into E. We have following representation for the covariance operator of u,

$$(\mathrm{Cov}(u))\,(x^*) = \int_E \langle x, x^*\rangle\, x\, \mathcal{P}_u(\mathrm{d}x) - \int_E \langle x, x^*\rangle\, x_1\, \mathcal{P}_u(\mathrm{d}x)$$

$$- \int_E \langle x_1, x^*\rangle\, x\, \mathcal{P}_u(\mathrm{d}x) + \langle x_1, x^*\rangle\, x_1,$$

where $x_1 = E\{u\}$ and P_u is the probability distribution of the random variable u. Let Q be any probability measure on E such that $\int_E (p(x))^2\, Q(\mathrm{d}x) < \infty\ \forall\, p \in \mathcal{P}_E$ and x_1 be any fixed element of E. Then the linear operator T: $E^* \to E$, given by

$$Tx^* = \int_E \langle x, x^*\rangle\, xQ(\mathrm{d}x) - \int_E \langle x, x^*\rangle\, x_1\, Q(\mathrm{d}x)$$

$$- \int_E \langle x_1, x^*\rangle\, xQ(\mathrm{d}x) + \langle x_1, x^*\rangle\, x_1$$

is obviously a covariance operator of an E-valued random variable u on the probability space $(E, \mathcal{B}_E, Q)$ with $u(x) = x\ \forall x\ E$ and $E\{u\} = x_1$.

Definition 2.3.

An E-valued stochastic process $\{\omega(t), 0 \le t \le T\}$ of Gaussian† random variables is called a Wiener process if

(a) $\omega(t) \in L^2(\Omega, E, P)$ for all $t \in [0, T]$, $\omega(0) = 0$ a.s.,
(b) $E\{\omega(t) - \omega(s)\} = 0$,
(c) $\mathrm{cov}\{\omega(t) - \omega(s)\} = (t-s)\, W$,

where W: $E^* \to E$ is a continuous linear operator defined by

$$W e_n^* = \lambda_n\, e_n \tag{9}$$

such that $\lambda_n \ge 0$ and $\sum_{n=1}^{\infty} \sqrt{\lambda_n} < \infty$.

(d) $\{\omega(t), 0 \le t \le T\}$ is a process of independent increments and almost all paths of $\omega(t)$ are continuous.
Note that W has the following form

† An E-valued random variable u is called Gaussian if $\langle u, e_n^*\rangle$ is a real Gaussian random variable for all n.

$$Wx^* = \sum_{n=1}^{\infty} \langle x^*, e_n \rangle \lambda_n e_n.$$

Since E^* is barreled and E is quasi-complete, therefore W is a nuclear operator [10].

Lemma 2.3.

Let $\{\omega(t),\ 0 \le t \le T\}$ be an E-valued Wiener process. Then
(a) There exists a sequence $\{\beta_n(t)\}$ of mutually independent real Wiener processes such that

(i) $E\{\beta_n(t) - \beta_n(s)\}^2 = (t-s)\lambda_n$,

(ii) $\sum_{n=1}^{\infty} \sqrt{\lambda_n} < \infty$,

(iii) $\omega(t) = \sum_{n=1}^{\infty} \beta_n(t) e_n$.

(b) for every $p \in \mathcal{P}_E$ there exists a positive constant C such that

$$E[p\{\omega(t) - \omega(s)\}^2] \le C(t-s) \tag{10}$$

for $t \ge s$.

Proof: (a) Define $\beta_n(t) = \langle \omega(t), e_n^* \rangle$. It is easy to see that $\{\beta_n(t)\}$ is the required sequence of mutually independent real Wiener processes.

(b) Define $\omega_n(t) = \sum_{m=1}^{n} \beta_m(t) e_m$. For any $p \in \mathcal{P}_E$, $p(\omega_n(t))$ converges to $p(\omega(t))$ a.s. and $p(\omega_n(t)) \le M_p \sum_{m=1}^{\infty} |\beta_m(t)|$. Since $\sum_{m=1}^{\infty} \sqrt{\lambda_m} < \infty$, we have $\sum_{m=t}^{\infty} |\beta_m(t)| \in L^2(\Omega, \mathbf{R}, P)$. Rest of the proof follows by a simple application of Lebesgue dominated convergence theorem.

We have the following converse of the lemma 2.3 (a).

Lemma 2.4. Let $\{\beta_n(t)\}$ be a sequence of real valued mutually independent Wiener processes such that $E\{\beta_n(t) - \beta_n(s)\}^2 = (t-s)\lambda_n$ and $\sum_{n=1}^{\infty} \sqrt{\lambda_n} < \infty$. Then the process $\omega(t) = \sum_{n=1}^{\infty} \beta_n(t)\, e_n$ is an E-valued Wiener process.

Proof. Since $\sum_{n=1}^{\infty} \sqrt{\lambda_n} < \infty$, $\omega(t) \in L^2(\Omega, E, P)$. Since $\langle \omega(t), e_n^* \rangle = \beta_n(t)$, $\{\omega(t)\}$ is a Gaussian process. It is easy to verify that $\{\omega(t)\}$ is an E-valued Wiener process with $\operatorname{cov}\{\omega(t)-\omega(s)\} = (t-s)W$, where $W e_n^* = \lambda_n e_n$.

3. Stochastic integration

Let F be a countably Hilbert space, i.e., F is a complete locally convex space which topology is given by countable family of compatible Hilbertain norms $||, ||_n$, $n \in N$ such that $||, ||_{n_1} \leq ||, ||_{n_2}$ for all $n_1, n_2 \in N$ with $n_1 < n_2$. Since F is complete, we have $F = \bigcap_n F_n$, where F_n be the completion of F under the norm $||, ||_n$. Let E be a locally convex space of §2.3 and $\{\omega(t), 0 \leq t \leq T$ be a Wiener process in E. Let $L(E, F)$ be the space of all continuous linear operators from E into F. By Banach Steinhaus theorem, $L(E, F)$ be a sequentially complete locally convex space under the topology of simple convergence [11]. Let $\mathcal{B}_t = \sigma(\omega(s), s \leq t)$. Define

$M(E, F) = \{\phi(.,.)$: ϕ is an $L(E, F)$-valued measurable process adapted to $\mathcal{B}_t\}$,

$M_r^0(E, F) = \{\phi(.,.) \in M(E, F)$: ϕ is a t-step function on $[0, T]$ and $\phi \in L^r([0, T] \times \Omega, L(E, F), \mu \times p)$,

$M_r(E, F) = \{\phi(.,.) \in M(E, F) : \phi \in L^r([0, T] \times \Omega, L(E, F), \mu \times P)\}$

where $r \geq 1$ is an integer. Define a locally convex topology on $M_r(E, F)$ by family of semi-norms given by

$$q(\phi) = q(\phi, p) = [\int_0^T E\{p(\phi(t))\}^r \, dt]^{1/r}$$

for some $p \in \mathcal{P}_{L(E,F)}$ It is easy to see that $M_2^0(E, F)$ is dense in $M_2(E, F)$.

Definition 3.1

Let $\phi \in M_2^0(E, F)$ such that

$$\phi(t, \omega) = \sum_{j=1}^{N} \phi(t_j, \omega) \chi^{(t)}_{[t_j, t_{j+1}]}, \tag{11}$$

where $0 = t_1 < t_2 < \ldots < t_N = T$. The stochastic integral of ϕ with respect to the Wiener process $\{\omega(t), 0 \leq t \leq T\}$ is an F-valued random variable

denoted by $\int_0^T \phi(t)\, \mathrm{d}\omega(t)$ and is defined by

$$\int_0^T \phi(t)\, \mathrm{d}\omega(t) = \sum_{j=1}^{N} \phi(t_j)\,\{\omega(t_{j+1})-\omega(t_j)\}. \tag{12}$$

Lemma 3.2

If $\phi \in M_2^0\,(E, F)$, then

(a) $$E\left\{\int_0^T \phi(t)\, \mathrm{d}\omega(t)\right\} = 0$$

(b) Given any $n \in N$ there exist a constant $C > 0$ and $q \in \mathcal{P}_{L(E,F)}$ such that

$$E\left\| \int_0^T \phi(t)\, \mathrm{d}\omega(t) \right\|_n^2 \leq C \int_0^T E\{q(\phi(t))\}^2\, \mathrm{d}t.$$

Proof: (a) is an easy consequence of the fact that $\{\omega(t), 0 \leq t \leq T\}$ is a process of independent increments. The proof of (b) follows from the fact that if $A|\in|L(E, F)$ then $A \in L(E, F_n)$ for every n. Define a map $I: M_2^0(E, F) \rightarrow L^2(\Omega, F, P)$, by

$$I(\phi) = \int_0^T \phi(t)\, \mathrm{d}\omega(t).$$

For lemma 3.2b, we conclude that I is a continuous linear map from $M_2(E, F)$ into $L^2(\Omega, F, P)$ defined on $M_2^0(E, F)$. Since $L^2(\Omega, F, P)$ is complete and $M_2^0(E, F)$ is dense in $M_2(E, F)$, map I has unique extension $\bar{I}$ as a continuous linear map from $M_2(E, F)$ into $L^2(\Omega, F, P)$. For any $\phi \in M_2(E, F)$, $\bar{I}(\phi)$ will still be denoted by

$$\int_0^T \phi(t)\, \mathrm{d}\omega(t)$$

and is called the stochastic integral of ϕ with respect to Wiener process $\{\omega(t), 0 \leq t \leq T\}$. From lemma 3:2, we have

Theorem 3.3

If $\phi \in M_2(E, F)$, then

(a) $$\int_0^T \phi(t)\, \mathrm{d}\omega(t) = \sum_{i=1}^{\infty} \int_0^T \phi(t)\, e_i\, \mathrm{d}\beta_i(t), \tag{13}$$

(b) $$E\left\{\int_0^T \phi(t)\, \mathrm{d}\omega(t)\right\} = 0, \tag{14}$$

(c) Given any $n \in N$ there exist a constant $C > 0$ and $q \in \mathcal{P}_{L(E,F)}$ such that

$$E\left\|\int_0^T \phi(t)\, \mathrm{d}\omega(t)\right\|_n^2 \leq C\int_0^T E\{q(\phi(t))\}^2\, \mathrm{d}t. \tag{15}$$

Proof. (a) If $\phi \in M_2^0(E, F)$, then equation (13) is obvious. For $\phi \in M_2(E, F)$, equation (13) holds by the usual limiting arguments.

(b) Let $\phi_n \in M_2^0(E, F)$ such that ϕ_n converges to ϕ in $M_2(E, F)$ and $I(\phi_n)$ converges to $I(\phi)$. By Holder's inequality, we have

$$\lim_{n\to\infty} E\int_0^T \phi_n(t)\, \mathrm{d}\omega(t) \doteq E\int_0^T \phi(t)\, \mathrm{d}\omega(t)$$

in F, which completes the proof.

(c) Let $\{\phi_m\}_{m\in N}$ be same as in (b). Then for any $n \in N$, we have

$$E\left\|\int_0^T \phi(t)\, \mathrm{d}\omega(t)\right\|_n^2 \leq E\left\{\left\|\int_0^T (\phi_m-\phi)(t)\, \mathrm{d}\omega(t)\right\|_n \left\|\int_0^T \phi(t)\, \mathrm{d}\omega(t)\right\|_n\right\} + E\left\{\left\|\int_0^T (\phi_m-\phi)(t)\, \mathrm{d}\omega(t)\right\|_n \left\|\int_0^T \phi_m(t)\, \mathrm{d}\omega(t)\right\|_n\right\}$$
$$+E\left\|\int_0^T \phi_m(t)\, \mathrm{d}\omega(t)\right\|_n^2$$

By Holder's inequality

$$\begin{aligned}E\left\|\int_0^T \phi(t)\, \mathrm{d}\omega(t)\right\|_n^2 &\leq \lim_{m\to\infty} E\left\|\int_0^T \phi_m(t)\, \mathrm{d}\omega(t)\right\|_n^2,\\ &\leq C\lim_{m\to\infty}\int_0^T E\{q(\phi_m(t))\}^2\, \mathrm{d}t,\\ &= C\int_0^T E\{q(\phi(t))\}^2\, \mathrm{d}t\end{aligned}$$

for some positive constant C and for some $q \in \mathcal{P}_{L(E,F)}$.
Analogous to (2) and (3) we have

Lemma 3.4.

Let A be a closed linear operator on F and $\phi \in M_2(E, F)$ such that
(i) for all $t \in [0, T]$ and for all i

$$\phi(t)\, e_i \in \mathcal{D}(A) \text{ a.s.},$$

(ii) for every $n \in N$,

$$\sum_{i=1}^{\infty} E \int_0^T \| A\phi(s) e_i \|_n \, ds < \infty.$$

Then

(a) $$\int_0^T \phi(s) \, d\omega(s) \in \mathcal{D}(A),$$

(b) $$A \int_0^T \phi(s) \, d\omega(s) = \int_0^T A\phi(s) \, d\omega(s).$$

Lemma 3.5.

Let $U(t, s)$ be an almost strong evolution operator and $\phi \in M_2(E, F)$. Then

$$y_t = \int_0^t U(t, s) \, \phi(s) \, d\omega(s)$$

is strong continuous in mean square, i.e. $y_t \in C([0, T], L^2(\Omega, F, P))$.

Proof: Define

$$y_t^i = \int_0^t U(t, s) \, \phi(s) e_i \, d\beta_i(s).$$

For any $n \in N$ and $\delta > 0$, we have

$$E \| y_{t+\delta}^i - y_t^i \|_n^2 \leq 2E \| \, (U(t+\delta, t) - I) y_t^i \, \|_n^2$$

$$+ C \int_t^{t+\delta} E(q_1(\phi(s))\}^2 \, ds$$

for some $q_1 \in \mathcal{P}_{L(E,F)}$ and for some constant $C > 0$. Thus we have

$$\lim_{\delta \to 0} E \| y_{t+\delta}^i - y_t^i \|_n^2 = 0.$$

Similarly, we have for every $n \in N$ and $\delta > 0$ sufficiently small

$$\lim_{\delta \to 0} E\|y_t^i - y_{t-\delta}^i\|_n^2 = 0.$$

Using mutual independence of Wiener processes $\beta_i(t)$, it is easy to see that

$$\sum_{i=1}^{N} \int_0^t U(t, s)\, \phi(s) e_i \, d\beta_i(s) \text{ converges to}$$

$$\int_0^T U(t, s)\, \phi(s)\, d\omega(s)$$

in $C([0, T], L^2(\Omega, F, P))$, which completes the proof of lemma. The stochastic version of Fubini's type of theorem is not difficult to prove. We have

Lemma 3.6

Let $\psi(t, s, \omega) : [0, T] \times [0, T] \times \Omega \to L(E, F)$ be such that for each $x \in E$, $\psi(t, s, \omega)x$ is measurable on $[0, T] \times [0, T] \times \Omega$ and $\psi(t, \cdot, \cdot)x$ and $\psi(\cdot, t, \cdot)x$ are measurable relative to $\mathcal{B}_t$ for almost all $t \in [0, T]$. Suppose that for every $q \in \mathcal{P}_{L(E,F)}$,

$$\int_0^T \int_0^T E\{q(\phi)\}^2 \, ds \, dt < \infty.$$

If we define

$$y_1(\omega) = \sum_{i=1}^{\infty} \int_0^T \left\{ \int_0^T \psi(t, s, \omega) e_i \, ds \right\} d\beta_i(t),$$

$$y_2(\omega) = \int_0^T \left\{ \sum_{i=1}^{\infty} \int_0^T \psi(t, s, \omega) e_i \, d\beta_i(t) \right\} ds.$$

Then $y_1 = y_2$ a.s. and $y_1, y_2 \in L^2(\Omega, F, P)$.
The following lemma is crucial in the proof of Ito's formula.

Lemma 3.7. Let G be a countably Hilbert space and $\{X(t), 0 \le t \le T\}$ be a $L(F, L(F, G))$-valued stochastic process which is adapted to $\mathcal{B}_t$ and is such that $X(t) \in L^2(\Omega, L(F, G)), P)$ for all $t \in [0, T]$. Let Y_0 be a $L(E, F)$-valued random variable measurable with respect to $\mathcal{B}_{t_0}(0 \le t_0 < T)$ and $Y_0 \in L^4(\Omega, L(E, F), P)$. Then

$$E\{X(Y_0\, \Delta\omega)\, [Y_0\, \Delta\omega]/\mathcal{B}_s\} = (t-s) \sum_{i=1}^{\infty} \lambda_i\, X(Y_0\, e_i)\, [Y_0\, e_i] \tag{16}$$

for all s, t with $t_0 \leq s \leq t$, where $X = X(s)$ and $\Delta\omega = \omega(t) - \omega(s)$. The proof of the following theorem is similar to the proof of Curtain and Falb [1] of Ito's formula in Hilbert space.

Theorem 3.8 (Ito's formula)

Let E, F and G are as before and $\{\omega(t), 0 \leq t \leq T\}$ be an E-valued Wiener process. Let $\xi(t)$ be a F-valued stochastic process with stochastic differential $\mathrm{d}\xi(t) = b(t)\,\mathrm{d}t + \sigma(t)\,\mathrm{d}\omega(t)$ and $g: F \to G$ be a continuous function such that

(a) g is twice weakly differentiable,†

(b) Dg and D^2g are bounded continuous functions from F into $L(F, G)$ and $L(F, L(F, G))$ respectively,

(c) $b(t)$ is a F-valued stochastic process adapted to $\mathcal{B}_t$ and $\sigma \in M_2(E, F)$ such that

$$\int_0^T q(b(s))\,\mathrm{d}s < \infty \text{ a.s. } \forall\, q \in \mathcal{P}_F$$

and $\qquad \sigma \in L^4([0, T] \times \Omega, L(E, F), \mu \times P)$.

Then the process $z(t) = g(\xi(t))$ has G-valued stochastic differential

$$\mathrm{d}z(t) = \{Dg(\xi(t))b(t) + \frac{1}{2} \sum_{i=1} \lambda_i D^2 g(\xi(t))$$

$$(\sigma(t)e_i)\ [\sigma(t)e_i]\}\,\mathrm{d}t + [Dg(\xi(t))\ (\sigma(t))]\,\mathrm{d}\omega(t). \tag{17}$$

4. Stochastic evolution equation

Let $\{A(t), 0 \leq t \leq T\}$ be generator of almost strong evolution operator $U(.,.)$ in a countably Hilbert space F and $\{\omega(t), 0 \leq t \leq T\}$ be a Wiener process in a locally convex space E. Consider the following stochastic evolution equation

$$\begin{aligned} \mathrm{d}u(t) &= A(t)\,u(t)\,\mathrm{d}t + f(t)\,\mathrm{d}t + \phi(t)\,\mathrm{d}\omega(t), \\ u(0) &= u_0 \end{aligned} \tag{18}$$

in F, where $\phi \in M_2(E, F)$, $u_0 \in F$ and $f \in L^2([0, T] \times \Omega, F, \mu \times P)$. By (18), we mean that

$$u(t) = u_0 + \int_0^t A(r)\,u(r)\,\mathrm{d}t + \int_0^t f(r)\,\mathrm{d}r + \int_0^t \phi(r)\,\mathrm{d}\omega(r).$$

†For definition and other details about weak differentiability in locally convex space we refer [7].

Definition 4.1.

A solution $u(t)$ of (18) is said to be strong solution if $u(t) \in \mathfrak{D}(A(t))$ a.s., $u(t) \in C([0, T], L^2(\Omega, F, P))$ and $u(t)$ satisfies (18) a.s. on $T \times \Omega$. The strong solution $u(t)$ is said to be unique solution if whenever $v(t)$ is another solution

$$P\left\{\sup_{0 \leq t \leq T} \|u(t) - v(t)\|_n \neq 0\right\} = 0$$

for every $n \in N$.

Theorem 4.1

Assume that

(i) $U(t, 0)u_0 \in \mathfrak{D}(A(t))$ for every t,

(ii) $U(t, s)\,\phi(s)e_i \in \mathfrak{D}(A(t))$ a.s. for all i and for almost all $t > s$, and

$$\sum_{i=1}^{\infty} \lambda_i E\left\{\int_0^t \|A(t)\,U(t, r)\,\phi(r)e_i\|_n^2\,\mathrm{d}r\right\} < \infty$$

for every n, $E\|A(t)\,U(t, r)\,\phi(r)e_i\|_n^2 \in L^1([0, T] \times [0, T])$ for every n and every i,

(iii) $U(t, s)\,f(s) \in \mathfrak{D}(A(t))$ a.s. for almost all t, and

$$\int_0^t \|A(t)\,U(t, r)\,f(r)\|_n\,\mathrm{d}s < \infty$$

a.s. for every $n \in N$.

Then (18) has a unique strong solution with continuous sample path given by

$$u(t) = U(t, 0)u_0 + \int_0^t U(t, r)\,\phi(r)\,\mathrm{d}\omega(r)$$

$$+ \int_0^t U(t, r)\,f(r)\,\mathrm{d}r. \tag{19}$$

Proof. Uniqueness of solution follows from the uniqueness of solution of deterministic homogeneous evolution equation $\dot{x}(t) = A(t)\,x(t)$. From lemma 3.5 and strong continuity of $U(t, s)$ in t we have

$$\int_0^t U(t, r)\,\phi(r)\,\mathrm{d}\omega(r) \in C([0, T], L^2(\Omega, F, P))$$

and

$$\int_0^t U(t, r) f(r) \, dr \in C([0, T], L^2(\Omega, F, P)).$$

Thus we have $u(t) \in C([0, T], L^2(\Omega, F, P))$. It remains to show that $u(t)$ give (19) satisfies (18). From lemma 3.4, we have

$$A(t) \int_0^t U(t, r) \phi(r) \, d\omega(r) = \int_0^t A(t) U(t, r) \phi(r) \, d\omega(t) \text{ a.s.}$$

From lemma (3.6) and equation (20),

$$\int_0^t A(r) \left\{ \int_0^r U(r, \alpha) \phi(\alpha) \, d\omega(\alpha) \right\} dr$$

$$= \int_0^t \left\{ \sum_{i=1}^{\infty} \int_0^r A(r) U(r, \alpha) \phi(\alpha) e_i \, d\beta_i(\alpha) \right\} dr,$$

$$= \sum_{i=1}^{\infty} \int_0^t \left\{ \int_\alpha^t A(r) U(r, \alpha) \phi(\alpha) e_i \, dr \right\} d\beta_i(\alpha),$$

$$= \sum_{i=1}^{\infty} \int_0^t \{U(t, \alpha) \phi(\alpha) e_i - \phi(\alpha) e_i\} \, d\beta_i(\alpha),$$

$$= \int_0^t U(t, \alpha) \phi(\alpha) \, d\omega(\alpha) - \int_0^t \phi(\alpha) \, d\omega(\alpha).$$

Similarly, we can prove

$$\int_0^t A(r) \left\{ \int_0^r U(r, \alpha) f(\alpha) \, d\alpha \right\} dr$$

$$= \int_0^t U(t, \alpha) f(\alpha) \, d\alpha - \int_0^t f(\alpha) \, d\alpha.$$

From (21) and (22), we see that $u(t)$ satisfies (18)

Example. (Stochastic transport equation).

Let $F = L^2[0, 1]$ and A be a linear operator on $L^2[0, 1]$ with domain $\mathfrak{D}(A) = \{f \in L^2[0, 1]$, f is absolutely continuous and $f' \in L^2[0, 1]$, $f(1)$ defined by

$$Af = f'$$

Obviously A is the generator of a strongly continuous semi-group on $L^2[0, 1]$ and $\sigma(A) = \phi$. For each $t \in [0, T]$, let $A(t) = A$. It is easy to see that the family $\{A(t), t \in [0, T]\}$ of linear operators on $L^2[0, 1]$ satisfies conditions (1), (2), (3), (4) of theorem 2.1 and hence a generator of an almost strong evolution operator.

Let $E = L^p[0, 1]$ (for any $2 \leq p < \infty$) with $\{e_n\}$ as its Haar basis, namely,

$$e_1(t) = 1$$

$$e_{2^k+1}(t) = \begin{cases} \sqrt{2^k}, & t \in \left[\dfrac{2l-z}{2^{k+1}}, \dfrac{2l-1}{2^{k+1}}\right] \\ -\sqrt{2^k}, & t \in \left[\dfrac{2l-1}{2^{k+1}}, \dfrac{2l}{2^{k+1}}\right] \\ 0, \text{ for other } t. \end{cases}$$

$$(l = 1, 2, \ldots, 2^k,\ k = 0, 1, 2, \ldots).$$

Consider the following stochastic process in $L^p[0, 1]$:

$$\omega(t) = \sum_{n=1}^{\infty} \beta_n(t) e_n,$$

where $\{\beta_n(t)\}$ is a sequence of mutually independent real Wiener processes such that $E\{\beta_n(t) - \beta_n(s)\}^2 = (1/n^4)\ (t-s)$.

By lemma 2.4, $\{\omega(t)\}$ is an E-valued Wiener process.

Define a bounded linear operator $\mathcal{B}: L^p[0, 1] \to L^2[0, 1]$ by

$$(B\psi)(x) = \int_x^1 \psi(y)\, dy.$$

Let $\phi(t) = \mathcal{B}$(non-random) for every $t \in [0, T]$. Obviously $\phi \in M_2(E, F)$. Let $\{T(t), t \geq 0\}$ be the strongly continuous semigroup generated by A and let $U(t, s) = T(t-s), t \geq s$. $\{U(t, s)\}$ is an almost strong evolution operator generated by $\{A(t)\}$. Note that the semi-group $\{T(t), t \geq 0\}$ is obviously given by

$$T(t)f = g, \quad g(s) = \begin{matrix} f(s+t) & \text{for } s+t \leq 1 \\ 0 & \text{for } s+t > 1. \end{matrix}$$

We would like to solve the following stochastic transport equation:

$$\begin{aligned} du(t) &= u'(t)\, dt + B\, d\omega(t), \\ u_0 &\in \mathfrak{D}(A). \end{aligned} \tag{23}$$

Since $u_0 \in \mathfrak{D}(A)$, $Be_n \in \mathfrak{D}(A)$ for all n and $||AU(t, r)\, Be_n||_{L^2} \leq$ consta (independent of n), conditions (i) and (ii) of theorem 4.1 are satisfied. Therefo by theorem 4.1, equation (23) has a unique solution given by

$$u(t) = T(t) + \sum_{n=1}^{\infty} \int_0^t T(t-s)\, Be_n\, d\beta_n\,.$$

Acknowledgement

The author wishes to thank the referee for his suggestions which have led to t improvement of the paper.

References

[1] Curtain R F and Falb P L, Ito's formula in infinite dimensions, *J. Math. Anal. Appl.* **31** (19 434–448
[2] Curtain R F, Stochastic evolution equations with general white noise disturbance, *J. Math. An Appl.* **60** (1977) 570–595
[3] Curtain R F and Prichard A J, Infinite dimensional linear system theory, Springer Verlag, N York (1978)
[4] Gelfand I M and Vilenkin N Y, Generalized functions, Academic Press: New York (1964), Vo
[5] Ichikawa A, Linear stochastic evolution equations in hilbert space, *J. Diff. Equations* **28** (19 266–277
[6] Ito K, Stochastic analysis in infinite dimensions, *Proc. Int. Conf. on stochastic analysis* Acade Press: New York (1978) 187–197
[7] Keller H H, Differential calculus in locally convex spaces, Springer Verlag, New York (19
[8] Kuo H H, Stochastic integrals in abstract Wiener space II: regularity properties, *Nagoya Math* **50** (1973) 89–116
[9] Nevieu J, Descrete parameter martingale, North Holland: Amsterdam (1975)
[10] Treves F, Topological vector space, distributions and Kernels, Academic Press: New York (19
[11] Yoshida K, Functional Analysis, Springer Verlag, New York (1965)
[12] Yoshida K, Time dependent evolution equations in locally convex spaces, *Math. Annalen* (1965) 83–86

Proc. Indian Acad. Sci. (Math. Sci.), Vol. 95, No. 2, December 1986, pp. 97–107.

Sufficiency and strong commutants in quantum probability theory

SUBHASH J BHATT

Department of Mathematics, Sardar Patel University, Vallabh Vidyanagar 388 120, India

MS received 3 June 1985; revised 20 April 1986

Abstract. A probability algebra $(\mathscr{A}, *, \omega)$ consisting of a *algebra $\mathscr{A}$ with a faithful state ω provides a framework for an unbounded noncommutative probability theory. A characterization of symmetric probability algebra is obtained in terms of an unbounded strong commutant of the left regular representation of $\mathscr{A}$. Existence of coarse-graining is established for states that are absolutely continuous or continuous in the induced topology. Sufficiency of a *subalgebra relative to a family of states is discussed in terms of noncommutative Radon-Nikodym derivatives (a form of Halmos-Savage theorem), and is applied to a couple of examples (including the canonical algebra of one degree of freedom for Heisenberg commutation relation) to obtain unbounded analogues of sufficiency results known in probability theory over a von Neumann algebra.

Keywords. Quantum probability; unbounded representations; commutants; conditional expectation; coarse-graining; sufficiency; Schrödinger representation.

1. Introduction

A probability algebra $(\mathscr{A}, *, \omega)$ consists of a unital linear associative algebra $\mathscr{A}$ over complex numbers, an involution $x \to x^*$ on $\mathscr{A}$ making $\mathscr{A}$ a *algebra and *a* state ω on $\mathscr{A}$ that is faithful in the sense that $\omega(x^*x) = 0$ implies that $x = 0$. An inner product is defined on $\mathscr{A}$ as $\langle x, y\rangle = \omega(y^*x)$ satisfying $\langle xy, z\rangle = \langle y, x^*z\rangle$. Let H be the Hilbert space completion of $\mathscr{A}$ to which ω can be uniquely extended as a continuous linear form. A hermitian representation [9] $(\pi, D(\pi), H)$ of $\mathscr{A}$ is defined as $\pi(x)y = xy$ with domain $D(\pi) = \mathscr{A}$. This gives an Op*-algebra $\pi(\mathscr{A})$ of operators, not necessarily bounded.

Taking this as a basis, an unbounded noncommutative probability theory is developed in [2] which also encompasses unbounded observables in contrast to a bounded noncommutative probability theory considered till then (e.g. [3], [7], [11], [12]). These notes are aimed at further clarifying this foundation and to improve some of the results in [2] and [8]. The contributions of these notes are briefly summarized below.

(a) Through a noncommutative Radon-Nikodym theorem, the commutants of $\pi(\mathscr{A})$ [5] naturally enters into this scheme. In [2], the role of the weak commutant $\pi(\mathscr{A})^c$ is discussed. Here we examine the strong commutant $\pi(\mathscr{A})^c_s$. This clarifies

the inter-relations among various mathematical structures arising within the theory, e.g., a characterization of strong commutant is obtained analogous to the one for weak commutant [2, Theorem 6]. This implies that the probability algebra $\mathscr{A}$ is symmetric [2, §6] iff $\pi(\mathscr{A})^c_s$ is a *algebra. This bisects a result in [2, Theorem 6] that $\pi(\mathscr{A})^c$ is a *algebra iff $\mathscr{A}$ is symmetric and π is self-adjoint. We also discuss the probability algebra of simple operators associated with a noncommutative regular probability gauge space [13]. This is a noncommutative analogue of an example in [2, §2].

(b) In the discussion on conditional expectation and coarse-graining we establish the existence of a coarse-graining for a state that is absolutely continuous. This improves [2, Theorem 3(3)] wherein this has been shown for a dominated state. We also discuss the coarse-graining for states that are not absolutely continuous. This requires characterizing the linear forms on $\mathscr{A}$ that are continuous with respect to the induced topology [2, §6]. Information measures and coarse-graining relative to two different subalgebras are compared.

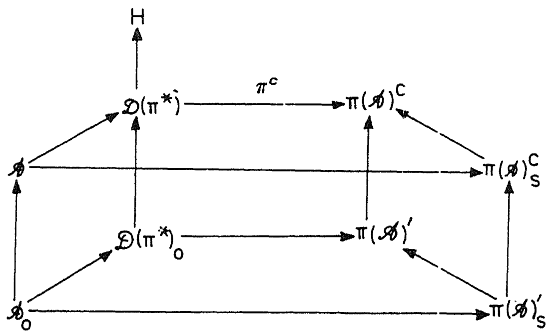
(c) Finally we discuss the sufficiency of *subalgebras of a probability algebra with respect to a family of state in terms of Radon-Nikodym derivatives. This is closely related with a noncommutative Halmos-Savage theorem on weak sufficiency (with a stronger notion of conditional expectation) obtained by Hiai *et al* [8]. Our sufficiency criterion leads to unbounded analogues of a number of results known in the context of probability theory over a von Neumann algebra. Let $\mathscr{A}$ be the canonical algebra of one degree of freedom for the Schrödinger representation of Heisenberg commutation relation. Let ω_β be the centred normal isotropic state [2, §7] of the parameter $\beta > 0$ on $\mathscr{A}$. Let G be the modular automorphism group associated with ω_β. Then an absolutely continuous state ν on $\mathscr{A}$ is G-invariant iff the centralizer of ω_β is sufficient for ν. This extends to the present framework, in the particular case of Schrödinger representation, a result for a von Neumann algebra with a faithful state obtained in [7, Theorem 2.2]. Likewise, we also obtain an analogue of [8, Example 1.3] for an abstract *algebra.

2. The strong commutants

Given a probability algebra $(\mathscr{A}, *, \omega)$ and the left regular representation $(\pi, D(\pi), H)$ the adjoint representation $(\pi^*, D(\pi^*), H)$ of $\mathscr{A}$ is $\pi^*(x) = \pi(x^*)^*|D(\pi^*)$ wherein the domain $D(\pi^*) = \cap\{D(\pi(x^*)^*|x \in \mathscr{A}\} \supset D(\pi)$ and $\pi \subset \pi^*$. The *right bounded part* of $\mathscr{A}$ is $\mathscr{A}_0 = \{x \in \mathscr{A}|\rho(x)$ is norm bounded$\}$, a subalgebra of $\mathscr{A}$. Here $\rho(x)y = yx$. The algebra $\mathscr{A}$ is *symmetric* if there exists an involution $y \to y^b$ on $\mathscr{A}$ such that $\omega(y^*x) = \omega(xy^b)$ $(x, y$ in $\mathscr{A})$. In this case, ρ is a ^{b}-antirepresentation of $\mathscr{A}$ and $\mathscr{A}_0$ is a ^{b}-subalgebra of $\mathscr{A}$.

Let $L(\mathscr{A})$ denote the set of all linear maps $\mathscr{A} \to H$. The *weak unbounded commutant of* π is $\pi(\mathscr{A})^c = \{C \in L(\mathscr{A})|\langle C\pi(x)y,z\rangle = \langle Cy,\pi(x^*)z\rangle$ for all x, y, z in

$\mathscr{A}\}$; where as the *weak bounded commutant* is $\pi(\mathscr{A})' = \{C \in \pi(\mathscr{A})^c | C$ is norm bounded$\}$. These hermitian linear subspaces are not algebras. On the other hand, the *strong unbounded commutant* $\pi(\mathscr{A})_s^c = \{C \in \pi(\mathscr{A})^c | C\mathscr{A} \subset \mathscr{A}\}$ and the *strong bounded commutant* $\pi(\mathscr{A})_s' = \{C \in \pi(\mathscr{A})_s^c | C$ is norm bounded$\}$ are algebras but not closed under adjoint. Let π^c: $D(\pi^*) \rightarrow \pi(\mathscr{A})^c$ be $\pi^c(y)x = \pi^*(x)y$. It is shown in [2, Theorem 8] that π^c is a linear isomorphism of $D(\pi^*)$ onto $\pi(\mathscr{A})^c$ which induces a linear isomorphism of $D(\pi^*)_0$ onto $\pi(\mathscr{A})'$ where $D(\pi^*)_0 = \{y \in D(\pi^*) | x \rightarrow \pi^c(y)x$ is norm bounded on $\mathscr{A}\}$; and that $\pi(\mathscr{A})^c$ is a *algebra (operator adjoint as the involution) iff π is self-adjoint and $\mathscr{A}$ is symmetric. In fact, it so turns out that π^c can be used to describe the strong commutants which in turn give a characterization of symmetric probability algebras. The following schematically exhibit the relationships of various commutants with other related structures.



2.1 *Theorem*

Let $(\mathscr{A}, *, \omega)$ be a probability algebra

(a) π^c is a bijective anti-isomorphism of $\mathscr{A}_0$ onto $\pi(\mathscr{A})_s^c$ which induces an anti-isomorphism of $\mathscr{A}$ onto $\pi(\mathscr{A})_s'$.

(b) $\mathscr{A}$ is symmetric iff $\pi(\mathscr{A})_s^c$ is a *algebra. In this case, π^c is a ^{b}-anti-isomorphism of $\mathscr{A}$ onto $\pi(\mathscr{A})_s^c$. Similarly, $\pi(\mathscr{A})_s'$ is a *algebra iff there exists an involution b on $\mathscr{A}_0$ such that $\omega(y^*x) = \omega(xy^b)$ (x, y in $\mathscr{A}_0$). Then π^c is a ^{b}-anti-isomorphism of $\mathscr{A}_0$ onto $\pi(\mathscr{A})_s'$.

Proof. (a) For x, y in $\mathscr{A}$, $\pi^c(y)x = xy \in \mathscr{A}$. Thus $\pi^c(y)\,\mathscr{A} \subset \mathscr{A}$, $\pi^c(y) \in \pi(\mathscr{A})_s^c$. For a $C \in \pi\,(\mathscr{A})_s^c$, $C1 \in \mathscr{A}$, say $C1 = y$. Then $\pi^c(y) = C$. It is easily checked that for y_1, y_2 in $\mathscr{A}$ and $x \in D(\pi)$, $\pi^c(y_1y_2)\,x = \pi^c(y_2)\,\pi^c(y_1)x$. Further, $y \in \mathscr{A}_0$ iff $\pi^c(y)$ is a bounded operator.

(b) Let $\pi(\mathscr{A})_s^c$ be an Op*-algebra. Then for a $C \in \pi(\mathscr{A})_s^c$, $C^* = C^\#|_{\mathscr{A}} \in \pi(\mathscr{A})_s^c$, $C^\#$ being the operator adjoint; and so $C^\#\mathscr{A} \subset \mathscr{A}$. Now given

$y \in \mathcal{A}$, $\pi^c(y) \in \pi(\mathcal{A})^c_s$, and so does $\pi^c(y)^*$. Define $y^b = \pi^c(y)^* 1$ $\in D(\pi) = \mathcal{A}$. Then for all $x \in \mathcal{A}$,

$$\omega(xy^b) = \langle xy^b, 1\rangle = \langle y^b, x^*\rangle = \langle \pi^c(y)^* 1, x^*\rangle = \langle 1, \pi^c(y)x^*\rangle$$
$$= \langle 1, \pi^*(x^*)y\rangle = \langle 1, x^*y\rangle = \langle x, y\rangle = \omega(y^*x).$$

Hence by [2, Lemma 7], $y \to y^b$ is an involution on $\mathcal{A}$ and $\mathcal{A}$ is symmetric. Conversely, if $\mathcal{A}$ is symmetric with an involution b such that $\omega(y^*x) = \omega(xy^b)$ $(x, y$ in $\mathcal{A})$; then for all x, y, z in $\mathcal{A}$,

$$\langle \pi^c(y^b)x, z\rangle = \langle xy^b, z\rangle = \langle x, zy\rangle = \langle \pi^c(y)^*x, z\rangle.$$

Hence $\pi^c(y)^* = \pi^c(y^b) \in \pi(\mathcal{A})^c_s$ in view of (a). Since $\pi^c(\mathcal{A}) = \pi(\mathcal{A})^c_s$, it follows that $\pi(\mathcal{A})^c_s$ is an Op*-algebra. Now it is clear that π^c is a ^{b}-anti-isomorphism. The remaining assertions can be analogously verified. This completes the proof.

We also note that $\pi(\mathcal{A})^c$ is closed under operator adjoint iff there exists a conjugate linear map $^b: D(\pi^*) \to D(\pi^*)$ such that $\langle x, y\rangle = \langle y^b, x\rangle$ for all $x \in \mathcal{A}$, $y \in D(\pi^*)$.

A state $F: \mathcal{A} \to \mathbb{C}$ is *absolutely continuous* (with respect to ω) $F < \omega$ if for any sequence (x_j) in $\mathcal{A}$, $\omega(x_j{}^*x_j) \to 0$ implies that $F(yx_j) \to 0$ for all $y \in \mathcal{A}$. F is *dominated* (by ω) $F < \omega$ if for some constant M, $F(x^*x) \leqslant M\omega(x^*x)$ $(x \in \mathcal{A})$. Clearly $F < \omega \Longrightarrow F < \omega$. A Radon-Nikodym theorem [2, Theorem 1] shows that $F < \omega$ iff there exists a $C \geqslant 0$ in $\pi(\mathcal{A})^c$ (called the *Radon Nikodym derivative* denoted by $\mathrm{d}F/\mathrm{d}\omega$). such that $F(x) = \omega(Cx)$ $(x \in \mathcal{A})$; and F is dominated if $(\mathrm{d}F/\mathrm{d}\omega) \in \pi(\mathcal{A})^{\cdot}$. It follows from the above that for an $F < \omega$, $(\mathrm{d}F/\mathrm{d}\omega) \in \pi(\mathcal{A})^c_s$ iff $F(x) = \omega(z^*x)$ for a unique $z \in \mathcal{A}$, and $(\mathrm{d}F/\mathrm{d}\omega) \in \pi(\mathcal{A})'_s$ iff $z \in \mathcal{A}_0$.

We end this discussion with a noncommutative analogue of an instructive example (the probability algebra of the classical probability space) considered in [2], and characterize the absolutely continuous states and the dominated states.

2.2 *Example.* Let $\Gamma = \{H, \mathcal{M}, m\}$ be a regular probability gauge space [13] viz. $\mathcal{M}$ is a von Neumann algebra on a Hilbert space H with a faithful normal tracial state m. (In fact, we can take m to be a faithful normal semifinite trace). An operator T on H is simple [14] if $T = \Sigma\,\lambda_i E_i$, a finite sum with λ_i's in $\mathbb{C}$ and E_i's mutually disjoint projections in $\mathcal{M}$. Simple operators in $\mathcal{M}$ forms a probability algebra $\mathcal{A}$ with a faithful state $\omega(T) = \Sigma\lambda_i m(E_i)$. For $1 \leqslant p \leqslant \infty$, let $L^p(\mathcal{M})$ be the noncommutative L^p-spaces associated with $\mathcal{M}$ as in [14]. A version of Radon-Nikodym theorem in [14] implies (since $\mathcal{A}$ is ultraweakly dense in $\mathcal{M}$) that if f is an ultraweakly continuous linear form on $\mathcal{A}$, then there exists a unique $A \in L^1(\mathcal{M})$ (denoted by $A = D_\omega(f)$) such that $f(X) = m(D_\omega(f)X)$ $(X \in \mathcal{A})$. Now, using the noncommutative L^p-L^q duality and the corresponding Holder's inequality (see e.g. [13]), it can be shown that $f < \omega$ iff $D_\omega(f) \in L^2(\mathcal{M})$ and $f < \omega$ iff $D_\omega(f) \in L^\infty(\mathcal{M})$.

3. Conditional expectation and coarse graining

Let $(\mathscr{A}, *, \omega)$ be a probability algebra. Let $\mathscr{B}$ be a *subalgebra of $\mathscr{A}$ containing 1. Let $\pi|_{\mathscr{B}}$ be the restriction of π to $\mathscr{B}$ with domain $D(\pi|_{\mathscr{B}}) = \mathscr{B}$ considered as a representation on $\overline{\mathscr{B}}$, the norm closure of $\mathscr{B}$ in H. The *conditional expectation* of x in $\mathscr{A}$ given $\mathscr{B}$ is $E(x|\mathscr{B}) \in D((\pi|\mathscr{B})^*) \subset \overline{\mathscr{B}}$ such that for all $y \in \mathscr{B}$, $\omega(yx) = \omega[(\pi|\mathscr{B})^*(y)\, E(x|\mathscr{B})]$. It is shown in [2, Theorem 2] that $E(x|\mathscr{B})$ exists for all $x \in \mathscr{A}$ and $E(x|\mathscr{B}) = P_{\mathscr{B}}x = (\mathrm{d}\omega_{x^*}/\mathrm{d}\omega|\mathscr{B})_1$ where $P_{\mathscr{B}}: H \to \overline{\mathscr{B}}$ is the projection and $\omega_x(y) = \omega(xy)$.

Let ν be a linear form (state) on $\mathscr{B}$ such that $\nu < \omega|\mathscr{B}$. A linear form (state) ν_c on $\mathscr{A}$ is a $(\mathscr{B}, \omega)$ *coarse-graining* of ν if $\nu_c|\mathscr{B} = \nu$, $\nu_c < \omega$ and $\nu_c(x) = \nu_c[E(x|\mathscr{B})]$ for all $x \in \mathscr{A}$. It is shown in [2, Theorem 3] that a $(\mathscr{B}, \omega)$ coarse-graining ν_c of ν exists iff ν has an absolutely continuous extension ν_1 to $\mathscr{A}$ such that $(\mathrm{d}\nu_1/\mathrm{d}\omega)\, 1 \in \overline{\mathscr{B}}$. Further, if $\mathscr{B}$ is *positivity preserving* (in the sense that given $x \in \mathscr{A}$, there exists a sequence (y_i) in $\mathscr{B}$ such that $y_i^* y_i \to P_{\mathscr{B}}(x^*x)$), and if ν is dominated, then ν_c exists. We show that under the same condition, even if ν is absolutely continuous, it admits a coarse-graining.

3.1 *Theorem*

Let $\mathscr{B}$ be a *subalgebra of a probability algebra $(\mathscr{A}, *, \omega)$.
(a) Let $\mathscr{B}$ be positivity preserving and $\nu: \mathscr{B} \to \mathbb{C}$ be a state such that $\nu < \omega|\mathscr{B}$. Then a $(\mathscr{B}, \omega)$ coarse-graining ν_c of ν exists.
(b) Let $\mathscr{B}$ be positivity preserving in the strong sense and $(\pi|\mathscr{B})^*$ be a hermitian representation of $\mathscr{B}$. Let $\nu: \mathscr{B} \to \mathbb{C}$ be a state on $\mathscr{B}$ that is continuous with respect to the left induced topology $t^1_{\mathscr{B}}$ on $\mathscr{B}$ due to $\mathscr{B}$. Then a 'coarse-graining' $\nu_c: \mathscr{A} \to \mathbb{C}$ of ν exists satisfying the following:

(i) $\nu_c|\mathscr{B} = \nu$,
(ii) $\nu_c(x) = \nu(E(x|\mathscr{B}))$ for all $x \in \mathscr{A}$
(iii) ν_c is continuous in the left induced topology $t^l_{\mathscr{A}}$ on $\mathscr{A}$ due to $\mathscr{A}$
(iv) ν_c is a state.

The *left-induced topology* $t^l_{\mathscr{B}}$ on a *subalgebra $\mathscr{B}$ is the locally convex topology defined by the seminorms $x \in \mathscr{B} \to \|\pi(y)x\|$ $(y \in \mathscr{B})$. This is a natural topology [9] one considers on the domain of an unbounded representation, and it coincides with the norm topology if each $\pi(y)$ is bounded. We say that $\mathscr{B}$ is *positivity preserving in the strong sense* provided given $x \in \mathscr{A}$, there exists a sequence (y_i) in $\mathscr{B}$ such that for each $y \in \mathscr{B}$, $\|\pi(y)y_i^* y_i - (\pi|\mathscr{B})^*(y)\; E(x^*x|\mathscr{B})\| \to 0$. Note that if $\pi(\mathscr{B})$ (and, in particular $\pi(\mathscr{A})$) is an algebra of bounded operators, then positivity preserving and strong positivity preserving are identical concepts. Our proof of (a) is a modification of that of [2, Theorem 3], whereas that of (b) requires determining the linear forms on $\mathscr{A}$ that are continuous in the induced topology.

Proof. (a) ν is $\|\cdot\|$ continuous on $\mathscr{B}$, hence admits a unique continuous linear extension $\nu\colon \overline{\mathscr{B}} \to \mathbb{C}$. Since $E(x|\mathscr{B}) \in D((\pi|\mathscr{B})^*) \subset \overline{\mathscr{B}}$, we can define a linear map $\nu_c\colon \mathscr{A} \to \mathbb{C}$, $\nu_c(x) = \nu(E(x|\mathscr{B}))$. To show that $\nu_c < \omega$, take a sequence (x_i) in $\mathscr{A}$ such that $\omega(x_i^* x_i) \to 0$, let $y \in \mathscr{A}$. Then $E(yx_i|\mathscr{B}) \in \overline{\mathscr{B}}$ for all i. For a fixed x_i, choose a sequence (z_n) in $\mathscr{B}$ such that $z_n \to E(yx_i|\mathscr{B})$ in norm. Then

$$\begin{aligned}
\nu_c(yx_i) &= \nu(E(yx_i|\mathscr{B}))\lim_n \nu(z_n) = \lim_n \omega\left(\frac{d\nu}{d\omega|\mathscr{B}} z_n\right) \\
&= \lim_n \left\langle z_n, \frac{d\nu}{d\omega|\mathscr{B}} 1\right\rangle = \left\langle P_{\mathscr{B}}(yx_i), \frac{d\nu}{d\omega|\mathscr{B}} 1\right\rangle \\
&= \left\langle yx_i, \frac{d\nu}{d\omega|\mathscr{B}} 1\right\rangle \quad \text{since} \quad \frac{d\nu}{d\omega|\mathscr{B}} \in D((\pi|\mathscr{B})^*) \subset \overline{\mathscr{B}} \\
&= \left\langle x_i, \pi^*(y^*)\frac{d\nu}{d\omega|\mathscr{B}} 1\right\rangle \text{ since } D((\pi|\mathscr{B})^*) \subset D(\pi^*) \\
&\leqslant \left\|\pi^*(y)\frac{d\nu}{d\omega|\mathscr{B}} 1\right\| \omega(x_i^* x_i)^{1/2} \to 0 \text{ as } i \to \infty.
\end{aligned}$$

The positivity preserving condition implies that ν_c is a state if ν is a state.

(b) Since $\nu\colon \mathscr{B} \to \mathbb{C}$ is continuous in $t^l_{\mathscr{B}}$, there exist a constant C and an $a \in \mathscr{B}$ such that

$$|\nu(x)| \leqslant C\|\pi(a)x\| \leqslant C\langle\pi(y)x, x\rangle \quad (y = 1 + a^*a).$$

Hence there exists a $z \in \overline{\mathscr{B}}$ such that $\nu(x) = \langle(\pi|\mathscr{B})^*(y)x, z\rangle$ for $x \in D((\pi|\mathscr{B})^*)$. ν_c on $\mathscr{A}$ is defined as

$$\nu_c(x) = \langle(\pi|\mathscr{B})^*(y)\, E(x|\mathscr{B}), z\rangle = \nu(E(x|\mathscr{B})).$$

Then for some sequence (z_n) in $\mathscr{B}$ with $\|z_n - z\| \to 0$,

$$\begin{aligned}
\nu_c(x) &= \lim_n \langle(\pi|\mathscr{B})^*(y)\, E(x|\mathscr{B}), z_n\rangle \\
&= \lim_n \omega(z_n^*(\pi|\mathscr{B})^*(y)\, E(x|\mathscr{B})) \\
&= \lim_n \omega((\pi|\mathscr{B})^*(z_n^* y)\, E(x|\mathscr{B})) = \lim_n \omega(z_n^* yx) \\
&= \lim_n \langle yx, z_n\rangle = \langle yx, z\rangle = \langle\pi^*(y)x, z\rangle.
\end{aligned}$$

Thus ν_c is linear, continuous in the induced topology $t^l_{\mathscr{A}}$, hence extends uniquely to $D(\pi^*)$. Also given $x \in \mathscr{A}$, for some sequence (y_n) in $\mathscr{B}$,

$$\nu_c(x^*x) = \lim_n \langle((\pi|\mathscr{B}))^*(y)y_n^* y_n, z\rangle = \lim_n \nu(y_n^* y_n) \geqslant 0$$

showing that ν_c is a state. This completes the proof. (It should be noted that the assumption of positivity or strongly positivity is used only to conclude that the extended functioned ν_c is positive if ν is positive.)

We compare the coarse-graining relative to two different subalgebras. Given a state ν on $\mathscr{A}$ that is absolutely continuous, the *information measure* $I_\omega(\nu)$ of ν is a number satisfying $I_\omega(\nu) \geqslant 0$ for all ν, $I_\omega(\omega) = 1$ and for any two absolutely continuous states ν_1 and ν_2 on $\mathscr{A}$ that are *mutually singular*

$$\left(\text{i.e. } \left\langle \frac{d\nu_1}{d\omega} 1, \frac{d\nu_2}{d\omega} 1 \right\rangle = 0\right),$$

$I_\omega(\nu_1+\nu_2) = I_\omega(\nu_1)+I_\omega(\nu_2)$. It is shown in [2, Theorem 4] that

$$I_\omega(\nu) = \left\| \frac{d\nu}{d\omega} 1 \right\|^2,$$

and for a state ν on $\mathscr{B}$, if a $(\mathscr{B}, \omega)$ coarse-graining ν_c of ν exists, then it is the unique absolutely continuous extension of ν with minimum information. This, together with Theorem 4 and Corollary on p. 4, both in [2], can be used to conclude the following. We omit the proof.

3.2 *Proposition*

Let $(\mathscr{A}, *, \omega)$ be a probability algebra. Let $\mathscr{B}_1$ and $\mathscr{B}_2$ be *subalgebras of $\mathscr{A}$ such that $1 \in \mathscr{B}_1 \subset \mathscr{B}_2$. Let $\nu < \omega|\mathscr{B}_2$ be a state on $\mathscr{B}_2$. Let $\nu_c^2 = (\mathscr{B}_2, \omega)$ coarse-graining of ν and $\nu_c^1 = (\mathscr{B}_1, \omega)$ coarse-graining of ν. The following are equivalent.

(a) $\nu_c^1 = \nu_c^2$,

(b) $I_\omega(\nu_c^1) = I_\omega(\nu_c^2)$,

(c) $\nu_c^1|\mathscr{B}_2 = \nu$,

(d) $\dfrac{d\nu_c^2}{d\omega} 1 \in \overline{\mathscr{B}}_1$.

If ν is faithful on $\mathscr{B}_2$, then any of the above is equivalent to

(e) $\mathscr{B}_1 = \mathscr{B}_2$.

4. Sufficiency

Hiai *et al* [7] investigated sufficiency of states in a von Neumann algebra. In [8], they analyzed sufficiency and weak sufficiency in an abstract *algebra. However, the condition expectation considered in these papers is more restrictive. In fact, it may not exist at all [11]; whereas the one considered in the present framework always exists [2]. (We also refer to [4] for the discussion of relative merits of an analogous approach to condition expectation in von Neumann algebras.) In the present context, we define sufficiency as follows:

Let $(\mathscr{A}, *, \omega)$ be a probability algebra. Let $S_{\mathscr{A}}$ be the collection of all linear functionals (in particular, state) on $\mathscr{A}$ that are absolutely continuous. Let $S \subset S_{\mathscr{A}}$

and $\mathscr{B} \subset \mathscr{A}$ be a *subalgebra containing 1. Then $\mathscr{B}$ is *sufficient* for S if $\varphi(x) = \varphi(E(x|\mathscr{B}))$ for all $x \in \mathscr{A}$, $\varphi \in S$.

4.1 *Theorem*

The following are equivalent

(a) $\mathscr{B}$ is sufficient for S

(b) $\frac{\mathrm{d}\varphi}{\mathrm{d}\omega} \mathscr{B} \subset \overline{\mathscr{B}}$ for all $\varphi \in S$

(c) $\frac{\mathrm{d}\varphi}{\mathrm{d}\omega} \mathscr{B} \subset D((\pi|\mathscr{B})^*)$ for all $\varphi \in S$

(d) $\frac{\mathrm{d}\varphi}{\mathrm{d}\omega} 1 \in D((\pi|\mathscr{B})^*)$ for all $\varphi \in S$

(e) $\left.\frac{\mathrm{d}\varphi}{\mathrm{d}\omega}\right|_{\mathscr{B}} = \frac{\mathrm{d}\varphi|\mathscr{B}}{\mathrm{d}\omega|\mathscr{B}}$ for all $\varphi \in S$.

This is closely related with a kind of noncommutative Halmos-Savage theorem [8, Theorem 1] on weak sufficiency. However, the proof is transparent in view of [2, Theorem 3]. The objective here is to apply it to a couple of examples exhibiting sufficiency phenomena analogous to those discussed in the context of probability theory over a von Neumann algebra.

(A) Let $\mathscr{A}$ be a semifinite von Neumann algebra with a faithful normal semifinite trace τ. Given a normal state φ on $\mathscr{A}$, there exists a positive operator $T_\varphi = \mathrm{d}\varphi/\mathrm{d}\tau$ such that $\varphi(A) = \tau(T_\varphi A)$ $(A \in \mathscr{A})$. Example 1.3 in [7] implies that for any family S of normal states, the von Neumann algebra $\mathscr{B}$ generated by $\{(\mathrm{d}\varphi/\mathrm{d}\tau) | \varphi \in S\}$ is sufficient for S. The following extends this to the unbounded case.

4.2 *Example*

Let $(\mathscr{A}, *, \omega)$ be a probability algebra for which π is self-adjoint. Let $\mathscr{A}$ be symmetric admitting another involution $y \to y^b$ such that $\omega(y^*x) = \omega(xy^b)$ $(x, y$ in $\mathscr{A})$. Let S be a family of absolutely continuous states on $\mathscr{A}$. For $\varphi \in S$, $(\mathrm{d}\varphi/\mathrm{d}\omega) \in \pi(\mathscr{A})^c = \pi(\mathscr{A})^c_s = \pi^c(\mathscr{A})$ since π is self-adjoint, and so $(\mathrm{d}\varphi/\mathrm{d}\omega) = \pi^c(y_\varphi)$ for a unique $y_\varphi \in \mathscr{A}$. Let $\mathscr{B}$ be the ^{b}-subalgebra of $\mathscr{A}$ generated by $\{y_\varphi | \varphi \in S\}$.

Assertion. $\mathscr{B}$ is sufficient for S.

Proof. For $x \in \mathscr{A}$, $\varphi \in S$, choose (z_n) in $\mathscr{B}$ such that $z_n \to E(x|\mathscr{B})$. Then

$$\varphi(E(x|\mathscr{B}) = \lim_n \varphi(z_n) = \lim_n \omega\left(\frac{\mathrm{d}\varphi}{\mathrm{d}\omega} z_n\right) = \lim_n \omega(\pi^c(y_\varphi) z_n)$$

$$= \lim_n \ \omega(\pi^*(z_n)y_\varphi) = \lim_n \ \langle y_\varphi, \ z_n^* \rangle$$

$$= \lim_n \ \langle z_n, \ y_\varphi^b \rangle = \langle E(x|\mathscr{B}), \ y_\varphi^b \rangle = \langle E(x|\mathscr{B}), \ (\pi|\mathscr{B}) y_\varphi^b \ 1 \rangle$$

$$= \langle (\pi|\mathscr{B}) \ (y_\varphi^b)^* E(x|\mathscr{B}), \ 1 \rangle = \omega((\pi|\mathscr{B}) \ (y_\varphi^b)^* E(x|\mathscr{B}))$$

$$= \omega((\pi(\mathscr{B})^* \ (y_\varphi^{b*}) \, E(x|\mathscr{B}))$$

$$= \omega(y_\varphi^{b*} x) = \langle x, \ y_\varphi^b \rangle = \langle y_\varphi, \ x^* \rangle = \langle \pi^*(x) y_\varphi, \ 1 \rangle$$

$$= \langle \pi^c(y_\varphi) x, \ 1 \rangle = \left\langle \frac{\mathrm{d}\varphi}{\mathrm{d}\omega} \cdot x, \ 1 \right\rangle = \omega \left(\frac{\mathrm{d}\varphi}{\mathrm{d}\omega} \, x \right).$$

$$= \varphi(x).$$

(B) Let σ_t^φ be the modular automorphism group associated with a faithful normal state φ on a von Neumann algebra $\mathscr{A}$. Let $I(\varphi)$ be the at of all σ_t^φ-invariant normal states on $\mathscr{A}$ and Z_φ be the centralizer of φ. It is shown in [7, Theorem 2.2] that a normal state $\psi \in I(\varphi)$ iff Z_φ is sufficient for $\{\varphi, \ \psi\}$. The following describes analogous phenomena for the canonical algebra of one degree of freedom for the Heisenberg commutation relation [2].

4.3 *Example*

Let $\mathscr{A}$ be the *algebra generated by hermitian elements p and q satisfying $pq - qp = -i \ 1$. The Schrödinger representation $(\pi_0, \ D(\pi_0), \ L^2(\mathbf{R}))$ of $\mathscr{A}$, with domain $D(\pi_0) = S(\mathbf{R})$ (Schwartz space) is defined as

$$\pi_0(p) = p_0, \ \pi_0(q) = q_0, \ p_0 f = -i \frac{df}{dt}, \ q_0 f \ (t) = tf(t) \ \ (f \in S(\mathbf{R})).$$

For each

$$\beta > 0, \ \omega_\beta(x) = (1 - e^{-\beta})^{-1} \sum_{n=0}^{\infty} e^{-\beta} \, \langle \pi_0(x) \, f_n, \ f_n \rangle \ \ (x \in \mathscr{A})$$

f_n being the normalized Hermite functions, be the centred normal isotropic state of β. Then $(\mathscr{A}, \ *, \ \omega)$ $(\omega = \omega_\beta)$ is a symmetric probability algebra (Note that, π_0 being self-adjoint, the considerations of example 4.1 can be applied to $\mathscr{A}$). For each $z \in \mathbb{C}$, $\Delta(z)p = \cosh(\beta z)p - i \sinh(\beta z)q$, $\Delta(z)q = i \sinh(\beta z)p + \cosh(\beta z)q$ define an automorphism group G: $t \in \mathbf{R} \to \Delta(it)$, called the modular automorphism group of $\mathscr{A}$. Now let $I(\omega) = \{\varphi\colon \mathscr{A} \to \mathbb{C}$ state$| \ \varphi < \omega, \ \varphi(\Delta(it)x) = \varphi(x)$ for all $x \in \mathscr{A}$, $t \in \mathbf{R}\}$ be the set of all absolutely continuous G-invariant states on $\mathscr{A}$. Let $Z_\omega = \{y \in \mathscr{A} \,|\, \omega(xy) = \omega(yx)$ for all $x \in \mathscr{A}\}$ be the centralizer of ω.

Assertions. (a) For $x \in \mathscr{A}$, $t \in \mathbf{R}$, $E(x|Z_\omega) = E(\Delta(it)x|Z_\omega)$.
(b) For any state $\nu < \omega$ on $\mathscr{A}$, $\nu \in I(\omega)$ iff Z_ω is sufficient for ν.

Proof. Observe that, using [2, §7], exactly as in [10, Lemma 15.8], $Z_\omega = \{y \in \mathscr{A} \mid \Delta(it)y = y \text{ for all } t \in \mathbf{R}\}$, the fixed point algebra of $\mathscr{A}$.
(a) For $x \in \mathscr{A}$ let $\omega_x\colon Z_\omega \to \mathbb{C}$ be $\omega_x(y) = \omega(xy)$. Then by [2, Theorem 2(3)],

$$\omega_x < \omega \mid Z_\omega,\ E(x \mid Z_\omega) = \frac{\mathrm{d}\omega_{x^*}}{\mathrm{d}\omega \mid Z_\omega}\ 1.$$

Now for all $y \in Z_\omega$, $t \in \mathbf{R}$,

$$\omega_{x^*}(y) = \omega_{x^*}(\Delta(it)y) = \langle \Delta(it)y,\ x\rangle = \langle y,\ \Delta(-it)x\rangle$$
$$= \omega_{\Delta(-it)x^*}(y).$$

Hence

$$E(x \mid Z_\omega) = \frac{\mathrm{d}\omega_{\Delta(-it)x^*}}{\mathrm{d}\omega \mid Z_\omega} 1 = E(\Delta(-it)x \mid Z_\omega).$$

(b) Let Z_ω be sufficient for ν. Then for all $x \in \mathscr{A}$, $\nu(x) = \nu(E(x \mid Z_\omega))$. By Theorem 4.1, $(\mathrm{d}\nu/\mathrm{d}\omega)\ 1 \in \bar{Z}_\omega$ (throughout, the closure and the orthogonal complements are with reference to Hilbert space completion of the inner product structure on $\mathscr{A}$ defined by ω). Choose a (y_n) in Z_ω such that $y_n \to (\mathrm{d}\nu/\mathrm{d}\omega)\ 1$. Then for all $x \in \mathscr{A}$, $t \in \mathbf{R}$, using [2, Theorem 20],

$$\nu(\Delta(it)x) = \left\langle \frac{\mathrm{d}\nu}{\mathrm{d}\omega}\ 1,\ (\Delta(it)x)^*\right\rangle = \left\langle \frac{\mathrm{d}\nu}{\mathrm{d}\omega}\ 1,\ \Delta(it)x^*\right\rangle$$
$$= \lim_n\ \langle y_n,\ \Delta(it)x^*\rangle = \lim_n\ \langle \Delta(-it)y_n,\ x^*\rangle$$
$$= \lim_n\ \langle y_n,\ x^*\rangle = \left\langle \frac{\mathrm{d}\nu}{\mathrm{d}\omega}\ 1,\ x^*\right\rangle = \omega\left(\frac{\mathrm{d}\nu}{\mathrm{d}\omega}\ x\right).$$

Thus $\nu \in I(\omega)$. Conversely, let $\nu \in I(\omega)$. Then for all $x \in \mathscr{A}$, $t \in \mathbf{R}$, $\nu(x) = \nu(\Delta(it)x)$. Hence

$$\left\langle \frac{\mathrm{d}\nu}{\mathrm{d}\omega}\ 1,\ x^*\right\rangle = \omega\left(\frac{\mathrm{d}\nu}{\mathrm{d}\omega}\ x\right) = \nu(x) = \nu(\Delta(it)x)$$
$$= \omega\left(\frac{\mathrm{d}\nu}{\mathrm{d}\omega}\ \Delta(it)x\right) = \left\langle \frac{\mathrm{d}\nu}{\mathrm{d}\omega}\ 1,\ \Delta(it)x^*\right\rangle.$$

Thus $(\mathrm{d}\nu/\mathrm{d}\omega)\ 1 \perp A$ where $A = \{x - \Delta(it)x \mid x \in \mathscr{A},\ t \in \mathbf{R}\}$. Now let $z \in \mathscr{A} \cap A^\perp$. Then for all $x \in \mathscr{A}$, $t \in \mathbf{R}$, $\langle z, x\rangle = \langle z, \Delta(it)x\rangle = \langle \Delta(-it)z, x\rangle$. Hence $z = \Delta(-it)z$, $z \in Z_\omega$. Thus $\mathscr{A} \cap A^\perp \subset Z_\omega$; and so $(\mathscr{A} \cap A^\perp)^- \subset \bar{Z}_\omega$. But $(\mathscr{A} \cap A^\perp)^- = (\mathscr{A} \cap A^\perp)^{\perp\perp} = (\mathscr{A}^\perp \vee A^{\perp\perp})^\perp = A^\perp$. Hence $A^\perp \subset \bar{Z}_\omega$. It follows that $(\mathrm{d}\nu/\mathrm{d}\omega)\ 1 \in \bar{Z}_\omega$; and by Theorem 4.1, Z_ω is sufficient for ν.

Remark. In general, it can be shown that if $\mathscr{B}$ is a *subalgebra of a probability algebra $\mathscr{A}$, then $\mathscr{B}$ is sufficient for $S = \{\omega_x \mid x \in \mathscr{B}\}$, $\omega_x(y) = \omega(xy)$ $(y \in \mathscr{A})$.

References

[1] Gudder S P, A Radon-Nikodym Theorem for *algebras, *Pacific J. Math.* **80** (1979) 141–149

[2] Gudder S P and Hudson R L, A non-commutative probability theory, *Trans. Am. Math. Soc.* **235** (1976) 1–41

[3] Gudder S P and Marchand J P, Probability theory in Von Neumann algebras, *J. Math. Phys.* **13** (1972) 799–801

[4] Gudder S P and Marchand J P, Conditional expectation on Von.Neumann algebras; a new approach, *Rep. Math. Phys.* **12** (1977) 317–324

[5] Gudder S P and Scrugg W, Unbounded representations of *algebras, *Pacific J. Math.* **70** (1977) 369–382

[6] Halmos P R and Savage L J, Applications of Radon-Nikodym theorem to the theory of sufficient statistics, *Ann. Math. Stat.* **20** (1949) 225–241

[7] Hiai P, Ohya M and Tsukada M, Sufficiency, KMS condition and relative entropy in Von Neumann algebras, *Pacific J. Math.* **96** (1981) 99–109

[8] Hiai P, Ohya M and Tsukada M, Sufficiency and relative in *algebras with applications in quantum system, *Pacific J. Math.* **107** (1983) 117–140

[9] Powers R T, Selfadjoint algebras of unbounded operators, *Commun. Math. Phys.* **21** (1971) 88–124

[10] Takesaki M, *Tomita's theory of modular Hilbert algebras and its applications*, Springer Verlag lecture notes in Math. No. 128 (1971)

[11] Takesaki M, Conditional expectations in Von Neumann algebras, *J. Funct. Anal.* **9** (1972) 306–327

[12] Umegaki H, Conditional expectation in an operator algebra, *Tokuku Math. J.* **8** (1956) 86–100

[13] Wilde I F, The free Fermion field as a Markov field, *J. Funct. Anal.* **15** (1974) 12–22

[14] Yeadon F J, Noncommutative L^p-spaces, *Math. Proc. Cambridge Philos. Soc.* **77** (1975) 91–102

Proc. Indian Acad. Sci. (Math. Sci.), Vol. 95, No. 2, December 1986, pp. 109–120.

On torsional loading in an axisymmetric micropolar elastic medium

RAJNEESH KUMAR and T K CHADHA

Department of Mathematics, Guru Nanak Dev University, Amritsar 143 005, India

MS received 21 December 1985; revised 28 October 1986

Abstract. Effects of torsional loading in an axisymmetric micropolar elastic half-space are studied. The components of microrotation displacement, force stress and couple stress are obtained for a half-space subjected to an arbitrary load produced by shearing stress. A special case of a particular type of twist has been discussed in detail for a specific model and the micropolar effects have been shown graphically.

Keywords. Torsional loading; axisymmetric micropolar elastic medium; microrotation; displacement; force stress; couple stress.

1. Introduction

Modern engineering structures are often made up of materials possessing an internal structure. Polycrystalline materials, materials with fibrous or coarse grain structure come in this category. Classical elasticity is inadequate to represent the behaviour of such materials. The analysis of such materials requires incorporating the theory of oriented media. Describing the behaviour of oriented media, the micropolar theory is one of these theories.

The basic equations of the linear micropolar theory of elasticity have been given by Kuvchinski and Aero [9], Palmov [14], Eringen and Suhubi [3].

The first and second axial-symmetric problems in micropolar elasticity have been investigated by Nowacki [11,12]. Khan and Dhaliwal [4] also discussed the axisymmetric problem for a half-space in micropolar elastic medium. Nowacki and Nowacki [13] studied the axial symmetric Lamb's problem in micropolar elasticity. The present authors have already discussed the axisymmetric problem, the plane problem and Lamb's plane problem in micropolar elastic half-space with strech [5, 6, 7]. Kumar *et al* [8] also discussed Lamb's plane problem in thermoelastic micropolar medium with stretch.

Eason [1] discussed the problem of torsional loading of an elastic half-space. Wave motion due to impulsive twist on the surface has been investigated by Sarkar [15]. Sengupta [16] studied the problem of torsional vibration of a semi-infinite elastic medium by a particular type of twist. Tiwari [17] discussed the problem of effect of couple stress in a semi-infinite elastic medium due to impulsive twist on the surface.

In this paper, we first obtain the general solution for the problem of torsional loading in an axisymmetric micropolar elastic half-space under the action of an arbitrary load on its boundary and then the case of a particular type of twist has been discussed in detail.

2. Formulation of the problem

We consider a homogeneous, isotropic micropolar elastic half-space. We take cylindrical polar coordinate (r, θ, z), where z is pointing into the medium. Due to symmetry about z-axis, all quantities are independent of θ. We take

$$\boldsymbol{\omega} = (\omega_r, 0, \omega_z), \quad \mathbf{u} = (0, u_\theta, 0) \tag{1}$$

and assume that on the plane boundary $z = 0$ of the half-space $z \geqslant 0$, the only non-vanishing stress is the shearing stress which is given as $t_{z\theta} = -f(r)$, where $f(r)$ is the prescribed function of r. Moreover the displacement, stress and couple stress components tend to zero as $z \to \infty$.

3. Basic equations, boundary conditions and solutions

The constitutive equations and field equations in the absence of body forces and body moments for static case for micropolar elastic medium are given by Nowacki [10]:

$$t_{ji} = \lambda u_{k,k}\,\delta_{ji} + (\mu-\alpha)\,(u_{i,j}+u_{j,i}) + 2\alpha(u_{i,j} - \varepsilon_{kji}\omega_k), \tag{2}$$

$$m_{ji} = \beta\omega_{k,k}\delta_{ji} + (\gamma+\varepsilon)\omega_{i,j} + (\gamma-\varepsilon)\,\omega_{j,i}, \tag{3}$$

$$(\mu+\alpha)\nabla^2\mathbf{u} + (\lambda+\mu-\alpha)\ \text{grad div}\ \mathbf{u} + 2\alpha\ \text{rot}\ \boldsymbol{\omega} = 0, \tag{4}$$

$$(\gamma+\varepsilon)\nabla^2\boldsymbol{\omega} + (\gamma+\beta-\varepsilon)\ \text{grad div}\ \boldsymbol{\omega} + 2\alpha\ \text{rot}\ \mathbf{u} - 4\alpha\boldsymbol{\omega} = 0, \tag{5}$$

where λ, μ, α, β, γ, ε are material constants, $\mathbf{u}$ the displacement vector, $\boldsymbol{\omega}$ the rotation vector, t_{ji} the force stress tensor and m_{ji} the couple stress tensor. Using (1), (4) and (5) reduce to

$$\left[(\gamma+\varepsilon)\left(\nabla^2 - \frac{1}{r^2}\right) - 4\alpha\right]\omega_r + (\beta+\gamma-\varepsilon)\frac{\partial e'}{\partial r} - 2\alpha\frac{\partial u_0}{\partial z} = 0, \tag{6}$$

$$[(\gamma+\varepsilon)\nabla^2 - 4\alpha]\omega_z + (\beta+\gamma-\varepsilon)\frac{\partial e'}{\partial z} + \frac{2\alpha}{r}\frac{\partial}{\partial r}(ru_\theta) = 0, \tag{7}$$

$$(\mu+\alpha)\left(\nabla^2 - \frac{1}{r^2}\right)u_\theta + 2\alpha\left(\frac{\partial\omega_r}{\partial z} - \frac{\partial\omega_z}{\partial r}\right) = 0, \tag{8}$$

where

$$e' = \frac{1}{r}\frac{\partial}{\partial r}(r\omega_r) + \frac{\partial\omega_z}{\partial z}, \quad \nabla^2 = \frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{\partial^2}{\partial z^2}. \tag{9}$$

Since, there acts a torsional loading on the plane boundary $z = 0$, mathematically the boundary conditions are

$$t_{z\theta} = -f(r), \quad m_{zz} = m_{zr} = 0, \text{ at } z = 0. \tag{10}$$

To solve the problem, we introduce potential functions Φ, Ψ and ν as

$$\omega_r = \frac{\partial \Phi}{\partial r} + \frac{\partial^2 \Psi}{\partial r \partial z}, \tag{11}$$

$$\omega_z = \frac{\partial \Phi}{\partial z} - \left(\nabla^2 - \frac{\partial^2}{\partial z^2}\right)\Psi, \tag{12}$$

$$u_\theta = \partial \nu / \partial r. \tag{13}$$

Substituting (11) to (13) in (6) to (8), we get

$$\frac{\partial}{\partial r}[(\beta+2\gamma)\nabla^2 - 4\alpha]\Phi + \frac{\partial^2}{\partial r \partial z}[\{(\gamma+\varepsilon)\nabla^2 - 4\alpha\}\Psi - 2\alpha\nu] = 0, \tag{14}$$

$$\frac{\partial}{\partial z}[(\beta+2\gamma)\nabla^2 - 4\alpha]\Phi - \left(\nabla^2 - \frac{\partial^2}{\partial z^2}\right)[\{(\gamma+\varepsilon)\nabla^2 - 4\alpha\}\Psi - 2\alpha\nu] = 0, \tag{15}$$

$$\frac{\partial}{\partial r}(\nabla^2 \nu) + \frac{2\alpha}{(\mu+\alpha)}\frac{\partial}{\partial r}(\nabla^2 \Psi) = 0. \tag{16}$$

We use Hankel transforms defined by

$$\left.\begin{aligned} \bar{g}(\zeta, z) &= \int_0^\infty g(r, z) r J_0(\zeta r)\, dr, \\ \tilde{g}(\zeta, z) &= \int_0^\infty g(r, z) r J_1(\zeta r)\, dr, \end{aligned}\right\} \tag{17}$$

Applying transforms on (14) to (16), we get

$$\begin{aligned}[](\beta+2\gamma)(D_z^2 - \zeta^2) - 4\alpha]\bar{\Phi} + D_z[\{(\gamma+\varepsilon) \\ (D_z^2 - \zeta^2) - 4\alpha\}\bar{\Psi} - 2\alpha\bar{\nu}] = 0,\end{aligned} \tag{18}$$

$$\begin{aligned} D_z[(\beta+2\gamma)(D_z^2 - \zeta^2) - 4\alpha]\bar{\Phi} + \zeta^2[\{(\gamma+\varepsilon) \\ (D_z^2 - \zeta^2) - 4\alpha\}\bar{\Psi} - 2\alpha\bar{\nu}] = 0,\end{aligned} \tag{19}$$

$$\bar{\nu} = -\frac{2\alpha}{(\mu+\alpha)}\bar{\Psi}, \tag{20}$$

where $D_z = \mathrm{d}/\mathrm{d}z$. From (18) to (20), we get

$$(D_z^2 - \zeta^2)(D_z^2 - \zeta_1^2)\bar{\Phi} = 0, \tag{21}$$

$$(D_z^2 - \zeta^2)(D_z^2 - \zeta_2^2)\bar{\Psi} = 0, \tag{22}$$

$$(D_z^2 - \zeta^2)\bar{\Phi} = -\frac{(\gamma+\varepsilon)}{(\beta+2\gamma)} D_z(D_z^2 - \zeta^2)\bar{\Psi}, \tag{23}$$

where

$$\left.\begin{aligned}\zeta_1^2 &= \zeta^2+m_1^2,\ \zeta_2^2 = \zeta^2+m_2^2,\\ m_1^2 &= 4\alpha/(\beta+2\gamma),\ m_2^2 = 4\alpha\mu/(\mu+\alpha)\ (\gamma+\varepsilon).\end{aligned}\right\}$$

Since displacements, stress and couple stress components tend to zero as $z \to$ Φ, Ψ and ν tend to zero as $z \to \infty$, the solutions of (21) and (22) may be writter

$$\bar{\Phi} = A\exp(-\zeta z)+B\exp(-\zeta_1 z),$$

$$\bar{\Psi} = C\exp(-\zeta z)+D\exp(-\zeta_2 z).$$

Substituting (25) and (26) in (23), we obtain

$$A = \frac{\mu}{(\mu+\alpha)}\ \zeta C.$$

Thus

$$\bar{\Phi} = \frac{\mu}{(\mu+\alpha)}\zeta C\ \exp(-\zeta z)+B\ \exp(-\zeta_1 z).$$

Taking the inverse transform of (26) and (28)

$$\Phi = \int_0^\infty \left[\frac{\mu}{\mu+\alpha}\zeta C\ \exp(-\zeta z)+B\ \exp(-\zeta_1 z)\right]\zeta J_0(\zeta r)\,\mathrm{d}\zeta,$$

$$\Psi = \int_0^\infty [C\ \exp(-\zeta z)+D\ \exp(-\zeta_2 z)]\,\zeta J_0(\zeta r)\,\mathrm{d}\zeta.$$

Substituting the values of Φ and Ψ in the boundary conditions (10), we get following system of equations

$$2\alpha\zeta B+\frac{2\alpha\mu}{(\mu+\alpha)}\ \zeta^2 C = \bar{f}(\zeta),$$

$$(2\gamma\zeta^2+4\alpha)B-\frac{2\alpha\gamma}{(\mu+\alpha)}\ \zeta^3 C-2\gamma\zeta^2\zeta_2 D = 0,$$

$$2\gamma\zeta_1 B-\frac{2\alpha\gamma}{(\mu+\alpha)}\ \zeta^2 C-\left\{2\gamma\zeta^2-\frac{4\alpha\mu}{(\mu+\alpha)}\right\}D = 0,$$

where

$$\bar{f}(\zeta) = \int_0^\infty f(r)rJ_1(\zeta r)\,\mathrm{d}r.$$

Solving (31) to (33), we get

$$\left.\begin{aligned}& B = \zeta F_1(\zeta),\ C = \frac{(\mu+\alpha)}{\alpha}F_2(\zeta),\ D = F_3(\zeta)\\ &\text{where}\\ & F_1(\zeta) = (\zeta^2-\zeta\zeta_2+\varepsilon_1)\ \frac{\bar{f}(\zeta)}{2(\mu+\alpha)\Delta},\end{aligned}\right\}$$

$$\left.\begin{aligned} F_2(\zeta) &= \left[\zeta^2-\zeta_1\zeta_2+\left(\frac{(2\mu+\alpha)}{\mu}\right)\varepsilon_1+\frac{2\alpha}{\gamma}\frac{\varepsilon_1}{\zeta^2}\right]\frac{\bar{f}(\zeta)}{2(\mu+\alpha)\Delta},\\ F_3(\zeta) &= \left[\zeta(\zeta_1-\zeta)-\frac{2\alpha}{\gamma}\right]\frac{\bar{f}(\zeta)}{2(\mu+\alpha)\Delta},\\ \varepsilon_1 &= 2\alpha\mu/\gamma(\mu+\alpha),\ \Delta = (\zeta^2+\varepsilon_1)^2-\frac{(\alpha\zeta+\mu\zeta_1)}{(\mu+\alpha)}\zeta^2\zeta_2. \end{aligned}\right\} \tag{35}$$

and

Thus, we get the solutions of Φ and Ψ as

$$\Phi = \int_0^\infty \zeta^2\left[\frac{\mu}{\alpha}F_2(\zeta)\exp(-\zeta z)+F_1(\zeta)\exp(-\zeta_1 z)\right]J_0(\zeta r)\,\mathrm{d}\zeta, \tag{36}$$

$$\Psi = \int_0^\infty \zeta\left[\frac{(\mu+\alpha)}{\alpha}F_2(\zeta)\exp(-\zeta z)+F_3(\zeta)\exp(-\zeta_2 z)\right]J_0(\zeta r)\,\mathrm{d}\zeta, \tag{37}$$

Now, we obtain the components of displacement, microrotation, force stress and couple stress as

$$u_\theta = 2\int_0^\infty \zeta^2\left[F_1(\zeta)\exp(-\zeta z)+\frac{\alpha}{(\mu+\alpha)}F_3(\zeta)\exp(-\zeta_2 z)\right]J_1(\zeta r)\,\mathrm{d}\zeta, \tag{38}$$

$$\omega_r = \int_0^\infty \zeta^3\left[F_2(\zeta)\exp(-\zeta z)+\frac{\zeta_2}{\zeta}F_3(\zeta)\right.$$
$$\left.\times\exp(-\zeta_2 z)-F_1(\zeta)\exp(-\zeta_1 z)\right]J_1(\zeta r)\,\mathrm{d}\zeta. \tag{39}$$

$$\omega_z = \int_0^\infty \zeta^3\left[F_2(\zeta)\exp(-\zeta z)-\frac{\zeta_1}{\zeta}F_1(\zeta)\right.$$
$$\left.\times\exp(-\zeta_1 z)+F_3(\zeta)\exp(-\zeta_2 z)\right]J_0(\zeta r)\,\mathrm{d}\zeta, \tag{40}$$

$$t_{z\theta} = -2\mu\int_0 \zeta^3\left[F_2(\zeta)\exp(-\zeta z)+\frac{\alpha}{\mu}F_1(\zeta)\exp(-\zeta_1 z)\right]J_1(\zeta r)\,\mathrm{d}\zeta, \tag{41}$$

$$m_{zz} = -2\gamma\int_0^\infty \zeta^4\left[F_2(\zeta)\exp(-\zeta z)-F_1(\zeta)\left(1+\frac{2\alpha}{\gamma}\frac{1}{\zeta^2}\right)\exp(-\zeta_1 z)\right.$$
$$\left.+\frac{\zeta_2}{\zeta}F_3(\zeta)\exp(-\zeta_2 z)\right]J_0(\zeta r)\,\mathrm{d}\zeta, \tag{42}$$

$$m_{zr} = -2\gamma\int_0^\infty \zeta^4\left[F_2(\zeta)\exp(-\zeta z)-\frac{\zeta_1}{\zeta}F_1(\zeta)\exp(-\zeta_1 z)+F_3(\zeta)\right.$$

$$\times\left(1+\frac{\varepsilon_1}{\zeta^2}\right)\exp(-\zeta_2 z)\Big]J_1(\zeta r)\,\mathrm{d}\zeta. \tag{43}$$

If, we let the micropolar constants tend to zero, we obtain the displacement and stress components for the case of classical elasticity as

$$u_\theta = \frac{1}{\mu}\int_0^\infty \tilde{f}(\zeta)\exp(-\zeta z)\, J_1(\zeta r)\,\mathrm{d}\zeta, \tag{44}$$

$$t_{z\theta} = -\int_0^\infty \tilde{f}(\zeta)\zeta\exp(-\zeta z)\, J_1(\zeta r)\,\mathrm{d}\zeta. \tag{45}$$

These results agree with the results of Sarkar [15] for a static case.

4. Particular case

We consider the twist of the type given by

$$f(r) = \frac{r}{4a^4}\exp(-r^2/4a^4) \tag{46}$$

where r is the distance of the point from the origin of the co-ordinate system.

Applying the transform (17) on (46), we get

$$\tilde{f}(\zeta) = \zeta\exp(-a^2\zeta^2). \tag{47}$$

Thus, from (38) to (43) and (47), we get

$$u_\theta = 2\int_0^\infty \zeta^3\exp(-a^2\zeta^2)\Big[F_1(\zeta)\exp(-\zeta z)+\frac{\alpha}{(\mu+\alpha)}F_3(\zeta)$$

$$\times\exp(-\zeta_2 z)\Big]J_1(\zeta r)\,\mathrm{d}\zeta, \tag{48}$$

$$\omega_r = \int_0^\infty \zeta^4\exp(-a^2\zeta^2)\Big[F_2(\zeta)\exp(-\zeta z)+\frac{\zeta_2}{\zeta}F_3(\zeta)$$

$$\times\exp(-\zeta_2 z)-F_1(\zeta)\exp(-\zeta_1 z)\Big]J_1(\zeta r)\,\mathrm{d}\zeta, \tag{49}$$

$$\omega_z = \int_0^\infty \zeta^4\exp(-a^2\zeta^2)\Big[F_2(\zeta)\exp(-\zeta z)-\frac{\zeta_1}{\zeta}F_1(\zeta)$$

$$\times\exp(-\zeta_1 z)+F_3(\zeta)\exp(-\zeta_2 z)\Big]J_0(\zeta r)\,\mathrm{d}\zeta, \tag{50}$$

$$t_{z\theta} = -2\mu\int_0^\infty \zeta^4\exp(-a^2\zeta^2)\Big[F_2(\zeta)\exp(-\zeta z)+\frac{\alpha}{\mu}F_1(\zeta)$$

$$\times \exp(-\zeta_1 z)\Big] J_1(\zeta r)\,d\zeta, \tag{51}$$

$$m_{zz} = -2\gamma \int_0^\infty \zeta^5 \exp(-a^2\zeta^2)\Big[F_2(\zeta)\exp(-\zeta z) - F_1(\zeta)\left(1+\frac{2\alpha}{\gamma}\frac{1}{\zeta^2}\right)$$

$$\times \exp(-\zeta_1 z) + \frac{\zeta_2}{\zeta} F_3(\zeta)\exp(-\zeta_2 z)\Big] J_0(\zeta r)\,d\zeta, \tag{52}$$

$$m_{zr} = -2\gamma \int_0^\infty \zeta^5 \exp(-a^2\zeta^2)\Big[F_2(\zeta)\exp(-\zeta z) - \frac{\zeta_1}{\zeta} F_1(\zeta)$$

$$\times \exp(-\zeta_1 z) + \left(1+\frac{\varepsilon_1}{\zeta^2}\right) F_3(\zeta)\exp(-\zeta_2 z)\Big] J_1(\zeta r)\,d\zeta. \tag{53}$$

By neglecting the micropolar effect, we get the corresponding results for the classical elasticity as

$$u_\theta = \frac{1}{\mu}\int_0^\infty \zeta \exp(-a^2\zeta^2)\exp(-\zeta z)\, J_1(\zeta r)\,d\zeta, \tag{54}$$

$$t_{z\theta} = -\int_0^\infty \zeta^2 \exp(-a^2\zeta^2)\exp(-\zeta z)\, J_1(\zeta r)\,d\zeta. \tag{55}$$

5. Approximate evaluation of the integrals

The integrals involved in the expressions (48) to (53) are difficult to evaluate analytically, but can be evaluated either numerically or approximately. We evaluate them approximately.

Assuming α, m_1^2 and m_2^2 to be very small as compared to unity we expand ζ_1, ζ_2, $1/\Delta$ in an infinite series to obtain

$$\left.\begin{aligned} \zeta_1 &= \zeta + \frac{m_1^2}{2\zeta} + O(m_1^4), \\ \zeta_2 &= \zeta + \frac{m_2^2}{2\zeta} + O(m_2^4), \\ \Delta &= \gamma\varepsilon_1 A_1 \zeta^2, \end{aligned}\right\} \tag{56}$$

where

$$A_1 = \left[\frac{2}{\gamma} - \frac{1}{(\gamma+\varepsilon)} - \frac{1}{(\beta+2\gamma)}\right].$$

Hence, from (48) to (56), we get

$$u_\theta = \frac{1}{(\mu+\alpha)} \int_0^\infty \zeta\left(1+\frac{2\alpha P_1}{A_1\zeta^2}\right) \exp\,[(-a^2\zeta+z)\zeta]J_1(\zeta r)\,\mathrm{d}\zeta, \tag{57}$$

$$\omega_r = \frac{\alpha}{\gamma(\mu+\alpha)A_1} \int_0^\infty (P_4+P_2\zeta z) \exp\,[(-a^2\zeta+z)\zeta]J_1(\zeta r)\,\mathrm{d}\zeta, \tag{58}$$

$$\omega_z = \frac{\alpha}{\gamma(\mu+\alpha)A_1} \int_0^\infty (P_3+P_2\zeta z) \exp\,[(-a^2\zeta+z)\zeta]J_0(\zeta r)\,\mathrm{d}\zeta, \tag{59}$$

$$t_{z\theta} = -\int_0^\infty \left(\zeta^2+\gamma\varepsilon_1\frac{P_1}{A_1}\right) \exp\,[(-a^2\zeta+z)\zeta]J_1(\zeta r)\,\mathrm{d}\zeta, \tag{60}$$

$$m_{zz} = -\frac{2\alpha P_2}{(\mu+\alpha)A_1}\, z \int_0^\infty \zeta^2 \exp\,[(-a^2\zeta+z)\zeta]J_0(\zeta r)\,\mathrm{d}\zeta, \tag{61}$$

$$m_{zr} = -\frac{2\alpha P_2}{(\mu+\alpha)A_1}\, z \int_0^\infty \zeta^2 \exp\,[(-a^2\zeta+z)\zeta]J_1(\zeta r)\,\mathrm{d}\zeta, \tag{62}$$

where

$$P_1 = \frac{1}{\gamma^2} - \frac{1}{(\gamma+\varepsilon)(\beta+2\gamma)}, \quad P_2 = \frac{1}{(\gamma+\varepsilon)} + \frac{1}{(\beta+2\gamma)} - \frac{2}{(\gamma+\varepsilon)(\beta+2\gamma)},$$

$$P_3 = \frac{1}{\gamma} - \frac{1}{(\beta+2\gamma)}, \quad P_4 = \frac{1}{\gamma} - \frac{1}{(\gamma+\varepsilon)}. \tag{63}$$

We expand $\exp(-a^2\zeta^2)$ in an infinite series occurring in the expressions (57) to (62). Assuming that $a\zeta$ is so small that its fourth order terms are negligible and using the results of Erdelyi [2], we get

$$u_\theta = \frac{r}{(\mu+\alpha)}\left[\frac{1}{\rho_1^3}+\frac{3a^2}{\rho_1^5}\left(1-\frac{5z^2}{\rho_1^2}\right)+2\alpha\frac{P_1}{A_1}\right.$$
$$\left.\times\left\{\frac{1}{\rho_1+z}-\frac{a^2}{\rho_1^3}+\frac{3a^4}{2\rho_1^5}\left(\frac{5z^2}{\rho_1^2}-1\right)\right\}\right], \tag{64}$$

$$\omega_r = \frac{\alpha}{\gamma(\mu+\alpha)}\frac{r}{A_1\rho_1}\left[P_4\left\{\frac{1}{\rho_1+z}-\frac{3a^2z}{\rho_1^4}+\frac{15a^4z}{2\rho_1^6}\left(\frac{7z^2}{\rho_1^2}-3\right)\right\}\right.$$
$$\left.+P_2\frac{z}{\rho_1^2}\left\{1+\frac{3a^2}{\rho_1^2}\left(1-\frac{5z^2}{\rho_1^2}\right)\right\}\right], \tag{65}$$

$$\omega_z = \frac{\alpha}{\gamma(\mu+\alpha)A_1\rho_1}\left[P_3\left\{1+\frac{a^2}{\rho_1^2}\left(1-\frac{3z^2}{\rho_1^2}\right\}+\frac{9a^4}{2\rho_1^4}\left(1+\frac{5z^2}{\rho_1^2}\left(\frac{3z^2}{\rho_1^2}-2\right)\right)\right\}\right.$$

$$+P_2\frac{z^2}{\rho_1^2}\left\{1+\frac{3a^2}{\rho_1^2}\left(3-\frac{5z^2}{\rho_1^2}\right)\right\}\Bigg], \tag{66}$$

$$t_{z\theta}=-\frac{r}{\rho_1}\left[\frac{3z}{\rho_1^4}+\frac{15a^2z}{\rho_1^6}\left(3-\frac{7z^2}{\rho_1^2}\right)+\gamma\varepsilon_1\frac{P_1}{A_1}\right.$$

$$\left.\times\left\{\frac{1}{\rho_1+z}-\frac{3a^2z}{\rho_1^4}+\frac{15a^4z}{2\rho_1^6}\left(\frac{7z^2}{\rho_1^2}-3\right)\right\}\right], \tag{67}$$

$$m_{zz}=\frac{2\alpha zP_2}{(\mu+\alpha)A_1}\frac{1}{\rho_1^3}\left[1-\frac{3z^2}{\rho_1^2}+\frac{9a^2}{\rho_1^2}\left\{1+\frac{5z^2}{\rho_1^2}\left(\frac{3z^2}{\rho_1^2}-2\right)\right\}\right], \tag{68}$$

$$m_{zr}=\frac{6\alpha}{(\mu+\alpha)}\frac{P_2}{A_1}\frac{z^2r}{\rho_1^5}\left[\frac{5a^2}{\rho_1^2}\left(\frac{7z^2}{\rho_1^2}-3\right)-1\right], \tag{69}$$

where

$$\rho_1^2=r^2+z^2.$$

If we neglect the micropolar effect, we get the results for the classical elasticity

$$u_\theta=\frac{r}{\mu\rho_1^3}\left[1+\frac{3a^2}{\rho_1^2}\left(1-\frac{5z^2}{\rho_1^2}\right)\right], \tag{70}$$

$$t_{z\theta}=-\frac{3rz}{\rho_1^5}\left[1+\frac{5a^2}{\rho_1^2}\left(3-\frac{7z^2}{\rho_1^2}\right)\right], \tag{71}$$

6. Numerical results and discussion

Couple stress, force stress, microrotation and displacement have been calculated in the plane $z=1$ for various values of γ, $\alpha=0{\cdot}01$, $\beta=0{\cdot}03$, $\varepsilon=0{\cdot}02$, $\mu=0{\cdot}3$ and $a=1$ for the range $0\leqslant r\leqslant 4{\cdot}0$.

It is observed (figure 1) that for $\gamma=0{\cdot}05$, $0{\cdot}1$, $0{\cdot}15$, m_{zz} decreases for $0\leqslant r\leqslant 0{\cdot}8$, increases for $0{\cdot}8\leqslant r\leqslant 2{\cdot}8$ and decreases as r increases further for $r\geqslant 2{\cdot}8$.

It is evident from figure 2 that for $\gamma=0{\cdot}05, 0{\cdot}1, 0{\cdot}15$, ω_z increases monotonically in the intervals $0\leqslant r\leqslant 0{\cdot}4$ and $0{\cdot}8\leqslant r\leqslant 1{\cdot}6$, whereas it decreases monotonically for $0{\cdot}4\leqslant r\leqslant 0{\cdot}8$ and $r\geqslant 1{\cdot}6$.

We notice that for $\gamma=0{\cdot}05$, $t_{z\theta}$ increases monotonically for $0\leqslant r\leqslant 0{\cdot}4$ and $1{\cdot}2\leqslant r\leqslant 2{\cdot}8$ whereas it decreases monotonically for $0{\cdot}4\leqslant r\leqslant 1{\cdot}2$ and $r\geqslant 2{\cdot}8$. For $\gamma=0{\cdot}1$, $t_{z\theta}$ decreases monotonically for $0\leqslant r\leqslant 0{\cdot}4$ and $r\geqslant 2{\cdot}0$ but increases monotonically for $0{\cdot}4\leqslant r\leqslant 2{\cdot}0$. For $\gamma=0{\cdot}15$, $t_{z\theta}$ decreases monotonically for $0\leqslant r\leqslant 0{\cdot}4$ and $r\geqslant 1{\cdot}6$ but increases monotonically for $0{\cdot}4\leqslant r\leqslant 1{\cdot}6$. It is also

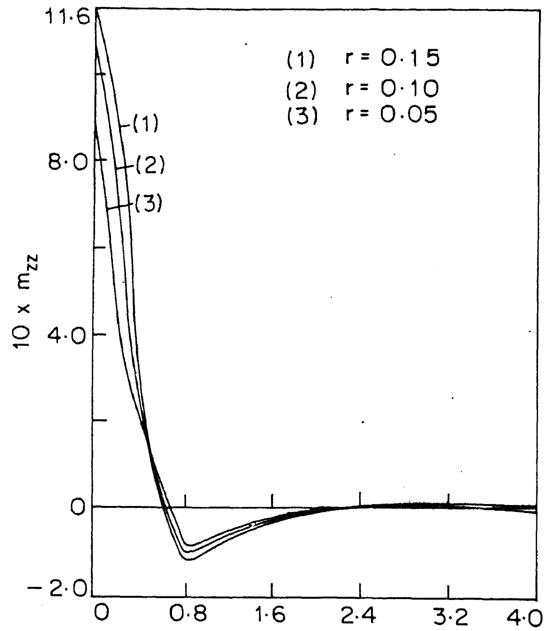


Figure 1. Couple stress m_{zz}.

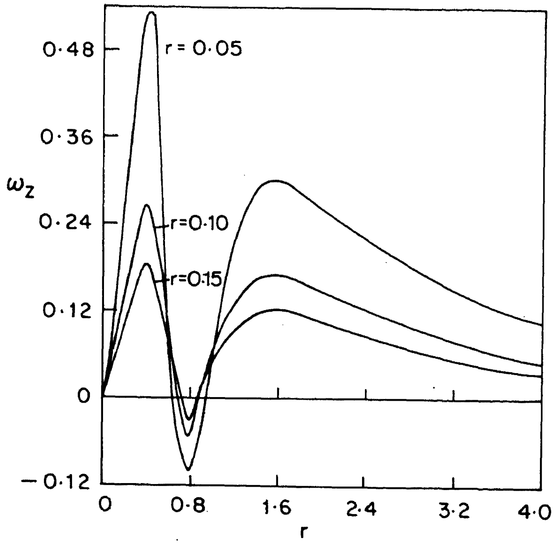


Figure 2. Microrotation ω_z.

observed that in the case of classical elasticity, $t_{z\theta}$ decreases monotonically for $0 \leqslant r \leqslant 0{\cdot}4$ and $r \geqslant 1{\cdot}6$ whereas it increases monotonically for $0{\cdot}4 \leqslant r \leqslant 1{\cdot}6$. These variations have been shown graphically in figure 3.

The variations of the displacement component u_θ have been shown graphically in figure 4. For $\gamma = 0{\cdot}05$, u_θ decreases monotonically for $0 \leqslant r \leqslant 0{\cdot}4$ and $r \geqslant 2{\cdot}4$

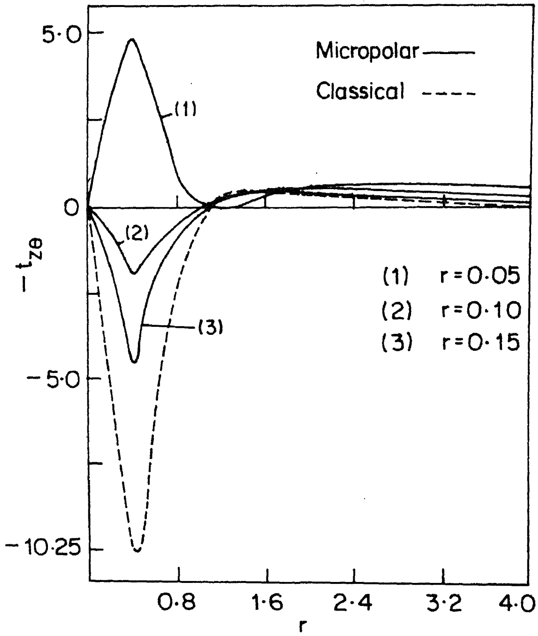


Figure 3. Force stress $t_{z\theta}$.

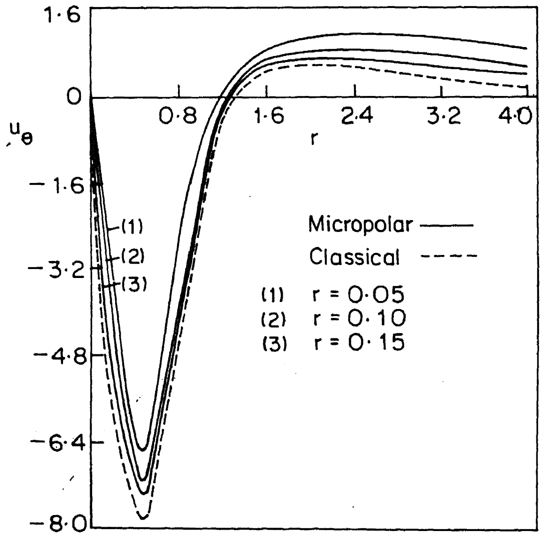


Figure 4. Displacement u_θ.

whereas it increases monotonically for $0{\cdot}4 \leqslant r \leqslant 2{\cdot}4$. For $\gamma = 0{\cdot}1$ and $0{\cdot}15$, u_θ decreases monotonically for $0 \leqslant r \leqslant 0{\cdot}4$ and $r \geqslant 2{\cdot}0$ whereas it increases monotonically for $0{\cdot}4 \leqslant r \leqslant 2{\cdot}0$. In the case of classical elasticity, u_θ decreases monotonically for $0 \leqslant r \leqslant 0{\cdot}4$, $r \geqslant 2{\cdot}0$ whereas it increases monotonically for $0{\cdot}4 \leqslant r \leqslant 2{\cdot}0$.

Thus, it is seen that force stress, couple stress, displacement and microrotation oscillate in certain range of r but all of them start decreasing beyond certain values of r.

Acknowledgement

The authors are thankful to Prof. Kehar Singh for help and suggestions.

References

[1] Eason G, On the torsional impulsive loading of an elastic half-space, *Q. J. Mech. Appl. Math.* **17** (1964) 279–292
[2] Erdelyi A, *Tables of integral transform* (New York: McGraw Hill) (1954) Vol. 2
[3] Eringen A C and Suhubi E S, Non-linear theory of simple microelastic solids, I and II, *Int. J. Eng. Sci.* **2** (1964) 189–203, 389–404
[4] Khan S M and Dhaliwal R S, Axisymmetric problem for half-space in the micropolar theory of elasticity, *J. Elast.* **7** (1977) 13–32
[5] Kumar R and Chadha T K, Axisymmetric problem in micropolar elastic half-space with stretch, *Int. J. Eng. Sci.* **23** (1985a) 1009–1029
[6] Kumar R and Chadha T K, Plane problem in micropolar elastic half-space with stretch, *Indian J. Pure Appl. Math.* **16** (1985b) 559–574
[7] Kumar R and Chadha T K, Lamb's plane problem in micropolar elastic half-space with stretch, *Bull. Acad. Pol. Sci. Ser. Sci. Tech.* **34** (1986) 59–72
[8] Kumar R, Chadha T K and Debnath L, Lamb's plane problem in thermoelastic micropolar medium with stretch, *Int. J. Math. Math. Sci.* **10** (1987) 187–198
[9] Kuvchinski E V and Aero E L, Continuum theory of asymmetric elasticity (in Russian) *Fiz. Trierd. Tiela* **5** (1963) 2592
[10] Nowacki W, Theory of micropolar elasticity, Int. Centre for Mechanical Sciences, Courses and Lectures, No. 25, (Berlin: Springer Verlag) (1970)
[11] Nowacki W, Axially symmetric problem in micropolar elasticity, *Bull. Acad. Pol. Sci. Ser. Sci. Tech.* **29** (1971) 317–325
[12] Nowacki W, Second axial symmetric problem in micropolar elasticity, *Bull. Acad. Pol. Sci. Ser. Sci. Tech.* **20** (1972) 517–522
[13] Nowacki W and Nowacki W K, The axial symmetric Lamb's problem in a semi-infinite micropolar elastic solid, *Proc. Vib. Probl. (Warsaw)* **10** (1969) 97
[14] Palmov M A, Fundamental equations of the theory of asymmetric elasticity (in Russian) *Prinkl. Math. Mekh.* **28** (1984) 401–408
[15] Sarkar G, Wave motion due to impulsive twist on the surface, *Pageoph* **65** (1966) 43–47
[16] Sengupta A K, Note on the torsional vibration of a semi-infinite elastic medium by a particular type of twist, *Bull. Calif. Math. Soc.* **56** (1964) 25–28
[17] Tiwari G, Effect of couple-stress in a semi-infinite elastic medium due to impulsive twist on the surface, *Pageoph* **91** (1971) 71–75

Proc. Indian Acad. Sci. (Math. Sci.), Vol. 95, No. 2, December 1986, pp. 121–125.

On a subclass of Bazilevic functions

T N SHANMUGAM

Department of Mathematics, Anna University, Madras 600 025, India

MS received 30 June 1986

Abstract. A new subclass of Bazilevic functions is defined and some of its properties have been studied.

Keywords. Bazilevic functions; subclass.

1. Introduction

Let S denote the class of functions $f(z) = z + \sum_{2}^{\infty} a_n z^n$ analytic, univalent in the open unit disk $U = \{z : |z| < 1\}$. For functions g and G analytic in U, we say that g is subordinate to G, denoted as $g \prec G$, if there exists a Schwarz function $w(z)$ analytic in U such that $g(z) = G(w(z))$, $z \in U$. In this paper a new class of functions $B(\alpha, h_1, h_2)$ is defined as follows:

Definition 1.1. Let h_1, h_2 be two convex univalent, analytic functions defined on U such that $h_1(U), h_2(U)$ lies on the right-half plane and $h_1(0) = h_2(0) = 1$. We say a function $f \in S$ belongs to the class $B(\alpha, h_1, h_2)$ if $zf'(z)/f^{1-\alpha}(z)\, g^{\alpha}(z) \prec h_2(z)$, $z \in U$ where $zg'(z)/g(z) \prec h_1(z)$, $z \in U$ with $\alpha \geqslant 0$.

For $\alpha = 0$ and $h_1(z) = h_2(z) = h(z)$, this class $B(\alpha, h_1, h_2)$ is the class $S^*(h)$, which is the same as the class $S_a^*(h)$ with $a = 1$ defined in [7]. Further if $h(z) = (1 + Az)/(1 + Bz)$ $(-1 \leqslant B < A \leqslant 1)$, $S^*(h)$ is the class $S^*(A, B)$ studied by Janowski [4]. For $\alpha = 1$ and for different choices of $h_1(z)$ and $h_2(z)$ the class $B(\alpha, h_1, h_2)$ generalizes various classes of functions introduced in [2, 5, 6, 12, 13]. In particular, for $h_2(z) = (1 + Az)/(1 + Bz)$ $(-1 \leqslant B < A \leqslant 1)$ and $h_1(z) = (1 + Cz)/(1 + Dz)$ $(-1 \leqslant D \leqslant C \leqslant 1)$, $B(\alpha, h_1, h_2)$ is the class $C[A, B; C, D]$ introduced by Silvia [12]. Further, if $B = D = -1$ and $A = 1 - 2\lambda$ $(0 \leqslant \lambda < 1)$ and $C = 1 - 2\rho$ $(0 \leqslant \rho \leqslant 1)$, then $B(\alpha, h_1, h_2)$ is the class. $B(\alpha, \lambda, \rho)$ was introduced by Gupta and Jain [3]. In this paper it is proved that the class $B(\alpha, h_1, h_2)$ is invariant under some integral operators.

To establish the main results of this paper, we need the following results:

THEOREM A [1]. Let $\beta, \gamma \in \mathbf{C}$ and h be a convex analytic univalent function in U with $h(0) = 1$ and $\operatorname{Re}(\beta h(z) + \nu) > 0$, $z \in U$. Let $p(z) = 1 + p_1 z + \ldots$ analytic in the unit disk. Then

$$p(z)+\frac{zp'(z)}{\beta p(z)+\nu} \prec h(z) \Longrightarrow p(z) \prec h(z), \quad z \in U.$$

A slight modification of the above result is as follows.

THEOREM B [8]. Let β, $\gamma \in \mathbf{C}$, $h(z)$ be a convex analytic univalent function in U with $h(0) = 1$ and $\mathrm{Re}(\beta q(z)+\gamma) > 0$, $z \in U$. If $p(z) = 1+p_1 z+\ldots$ is analytic in U then

$$p(z)+\frac{zp'(z)}{\beta q(z)+\gamma} \prec h(z) \Longrightarrow p(z) \prec h(z), \quad z \in U.$$

Let

$$f(z) = z+\sum_{2}^{\infty} a_n z^n, \; g(z) = z+\sum_{2}^{\infty} b_n z^n$$

then the Hadamards convolution product of f and g denoted by $(f^*g)\ (z)$ is

$$(f^*g)\ (z) = z+\sum_{2}^{\infty} a_n b_n z^n.$$

THEOREM C [10]. Let ϕ be a convex analytic function and $\phi(0) = 0$ and g is a star-like function in U. Then for F analytic in U with $F(0) = 1$, $[\phi^* Fg/\phi^* g]\ (U)$ is contained in the convex hall of $F(U)$.

2. Main theorems

THEOREM 2.1. Let $f \in S^*(h)$. Then for $\alpha > 0$ $\mathrm{Re}(c) > 0$, $F(z)$ defined as

$$F(z) = \left\{\frac{\alpha+c}{z^c}\int_0^z t^{c-1} f^\alpha(t)\ \mathrm{d}t\right\}^{1/\alpha}$$

is also an element of $S^*(h)$.

Proof. This follows easily by the application of Theorem A.

THEOREM 2.2. Let ϕ be a convex analytic function, then for $f \in S^*(h)$, $\phi * f \in S^*(h)$.

Proof. We have

$$\frac{z(\phi^* f)'(z)}{(\phi^* f)\ (z)} = \frac{(\phi^* zf')(z)}{(\phi^* f)\ (z)} = \frac{\left(\phi^* \frac{zf'}{f}\cdot f\right)(z)}{(\phi^* f)\ (z)}.$$

Since $f \in S^*(h)$ we have $zf'(z)/f(z) \prec h(z)$ in U; also ϕ is a convex function. Now an application of Theorem C yields that

$$\frac{z(\phi^* f)'(z)}{(\phi^* f)\ (z)} \prec h(z),$$

which establishes the theorem.

COROLLARY. If $f \in S^*(h)$, then

$$F_1(z) = \frac{c+1}{z^c}\int_0^z t^{c-1} f(t)\ \mathrm{d}t, \ \mathrm{Re}(c) > 0$$

$$F_2(z) = \int_0^z \frac{f(h)-f(xh)}{h-xh}\ \mathrm{d}h, \ |x| \leqslant 1, \ x \neq 1$$

also belongs to $S^*(h)$.

Proof. This is immediate from the theorem since F_1 and F_2 are convolutions of f with ϕ_1 and ϕ_2 being two convex functions respectively, where

$$\phi_1(z) = \sum_1^\infty \left(\frac{c+1}{c+n}\right) z^n, \ \phi_2(z) = \sum_1^\infty \frac{1-x^n}{(1-x)^n} z^n.$$

Remark. For the choice $h(z) = (1+Az)/(1+Bz)$, $-1 \leqslant A < B \leqslant 1$ in the above theorem and corollary we deduce theorem 8 and its corollary in [11].

THEOREM 2.3. Let ϕ be a convex function in U with $\phi(0) = 0$. Then $\phi^* f \in B(1, h_1, h_2)$ whenever $f \in B(1, h_1, h_2)$.

Proof. Now consider,

$$\frac{z(\phi^* f)'(z)}{(\phi^* g)(z)} = \frac{(\phi^* zf')(z)}{(\phi^* g)(z)} = \frac{\left(\phi^* \frac{zf'}{g}, g\right)(z)}{(\phi^* g)(z)}.$$

Since $f \in B(1, h_1, h_2)$ there exists a $g \in S^*(h_1)$ such that $zf'(z)/g(z) \prec h_2(z)$ in U. Hence an application of Theorem C yields the result.

COROLLARY. If $f \in B(1, h_1, h_2)$ then

$$F_1(z) = \frac{c+1}{z^c}\int_0^z t^{c-1} f(t)\ \mathrm{d}t, \ \mathrm{Re}(c) > 0$$

$$F_2(z) = \int_0^z \frac{f(h)-f(xh)}{h-xh}\ \mathrm{d}h, \ |x| \leqslant 1, \ x \neq 1$$

also belongs to $B(1, h_1, h_2)$.

Remark. When $h_1(z) = (1+Cz)/(1+Dz)$ $(-1 \leqslant D < C \leqslant 1)$ and $h_2(z) = (1+Az)/(1+Bz)$ $(-1 \leqslant B \leqslant A \leqslant 1)$, the above theorem reduces to theorem 4 in [12].

THEOREM 2.4. Let $\alpha > 0$, $\mathrm{Re}(c) > 0$ and let $f\ B(\alpha, h_1, h_2)$. Then $F(z)$ is defined by

$$F(z) = \left\{\frac{\alpha+c}{z^c}\int_0^z t^{c-1} f^\alpha(t)\ \mathrm{d}t\right\}^{1/\alpha} \tag{1}$$

is also belongs to $B(\alpha, h_1, h_2)$.

Proof. Since f is a Bazilevic function of type $(\alpha, 0)$ by a result due to Ruscheweyh [9], $F(z)$ defined by [1] is also a Bazilevic function of type $(\alpha, 0)$ and hence is univalent in U. Thus $F(z) \neq 0$ in $U-\{0\}$. Now, (1) on differentiation yields:

$$\frac{1}{z^\alpha}\{\alpha z F^{\alpha-1}(z)+CF^\alpha(z)\} = \frac{(\alpha+c)f^\alpha(z)}{z^\alpha}.$$

Since $f \in B(\alpha, h_1, h_2)$, $\exists\, g \in S^*(h_1)$ such that

$$zf'(z)/f^{1-\alpha}(z)g^\alpha(z) \prec h_2(z).$$

Let

$$p(z) = zF'(z)/F^{1-\alpha}(z)G^\alpha(z),$$

where

$$G(z) = \left\{\frac{C+\alpha}{z^C}\int_0^z t^{C-1}\, g^\alpha(t)\, \mathrm{d}t\right\}^{1/\alpha}.$$

Simplifying we get

$$\frac{zp'(z)}{\alpha\dfrac{zG'(z)}{G(z)}+C}+p(z) = \frac{zf'(z)}{f^{1-\alpha}(z)g^\alpha(z)} \prec h_2(z).$$

Also by Theorem 2.1, $G \in S^*(h_1)$. Now an application of Theorem B gives that $p(z) \prec h_2(z)$ when $\operatorname{Re}(c) > 0$, which means F is also in $B(\alpha, h_1, h_2)$ under the stated conditions of the theorem.

Remark 1. If

$$h_1(z) = \frac{1+(1-2\lambda)z}{1-z}$$

and

$$h_2(z) = \frac{1+(1-2\rho)z}{1-z}$$

in the above theorem, we get theorem 1 of [2].

Remark 2. If $\alpha = 1$ and $h_1(z) = h_2(z) = (1+Az)/(1+Bz)$ $(-1 \leqslant B < A \leqslant 1)$, in the above theorem we get a theorem due to Mehrok [6] as a particular case. Further if $A = 1$, $B = -1$ we get theorem 2 in [13].

Acknowledgement

The author thanks Prof. K S Padmanabhan for helpful suggestions during the preparation of this paper.

References

[1] Eenignburg P, Miller S S, Mocanu P T and Reade M O, On a Briot-Boquet differential subordinations, general inequalities 3 (Berlin: Birkause Verlag) Box 1, (1983) 339–348
[2] Goel R M and Mehrok B S, Some invariance properties of a subclass of close-to-convex functions, *Indian J. Pure Appl. Math.* **12** (1981) 1240–1249
[3] Gupta V P and Jain P K, A note on a class of Bazilevic function, *Tamkang J. Math.* **7** (1976) 117–119
[4] Janowski W, Some extremal problems for certain families of analytic functions I, *Am. Polon. Math.* **28** (1973) 297–326
[5] Kaplan W, Close-to-convex Schlicht-functions, *Mich. Math. J.* **1** (1952) 169–185
[6] Mehrok B S, A subclass of close-to-convex function, *Bull. Inst. Math. Acad. Sin.* **16** (1982) 389–398
[7] Padmanabhan K S and Parvatham R, Some applications of differential subordinations, *Bull. Aust. Math. Soc.* **32** (1985) 321–330
[8] Padmanabhan K S, Parvatham R and Shanmugam T N, α-convexity and α-close-to-convexity preserving integral operators, (submitted for publication)
[9] Ruscheweyh St, Eine invarianzeigenschaft der Bazilevic funktionen, *Math. Z.* **134** (1973) 215–219
[10] Ruscheweyh St and Shiel-Small T, Hadamard products of Schilicht functions and the polya-Schoenburg conjecture, *Comm. Math. Helv.* **48** (1983) 119–135
[11] Silverman H and Silvia E M, Subclasses of starlike functions to subordinates to convex functions, *Can. J. Math.* **37** (1985) 48–61
[12] Silvia E M, Subclasses of close-to-convex functions, *Int. J. Math. Math. Sci.* **6** (1983) 449–458
[13] Singh R, On Bazilevic functions, *Proc. Am. Math. Soc.* **38** (1973) 261–271

Proc. Indian Acad. Sci. (Math. Sci.), Vol. 95, No. 2, December 1986, pp. 127–132.

Computing the number of ways of representing primes by a norm form

RAJAT TANDON

School of Mathematics, University of Hyderabad, Hyderabad 500 134, India

MS received 11 September 1986; revised 12 February 1987

Abstract. Formulae for the number of different integral solutions of $a^2+b^2+c^2+d^2+ac+bd = p$ are given where p is a prime and the solution satisfies certain natural congruence conditions. Similar formulae are given for the case of the quadratic form $a^2+b^2+2c^2+2d^2+ac+bd$.

Keywords. Integral solutions; natural congruence.

The purpose of this note is to show how the results in [6] may be extended to the norm forms of other definite quaternion algebras ramifying at a single prime. The crucial extra information required is provided by [5]. The character formulas given in [5] enable us to "match" the representations of the idele group of the quaternion algebra and representations of $GL(2, \mathbf{A}_\mathbf{Q})$ with the same L-functions. The method is exemplified by considering definite quaternion algebras ramifying only at 3 and at 7. In determining the various representations, this paper incorporates an idea mentioned in [6] but not used. The notation used is of that in [6].

We will describe our quaternion algebra D in standard form. D is generated as a vector space over $\mathbf{Q}$ by $1, i, j, ij$ where $j^2 = -1$, $i^2 = -3$ and $ij = -ji$. This has one maximal order $\mathbf{Z}+\mathbf{Z}_j+\mathbf{Z}[\frac{1}{2}(1+i)]+\mathbf{Z}[\frac{1}{2}(1+i)]j$ (see [5]). The class number is 1. The order has twelve units $\{\pm 1, \pm j, \frac{1}{2}(\pm 1+i), \frac{1}{2}(\pm j \pm ij)\} = U$. The norm of any element of the form $a+bj+c[(1+i)/2]+[d(1+i)j]/2$ is $a^2+b^2+c^2+d^2+ac+bd$, denoted by $N(a,b,c,d)$.

Suppose p is an odd prime. Then it has been known for a long time that the number of integral solutions of $N(a,b,c,d) = p$ is $12(p+1)$ (see [2]).

We denote by D_A the adele ring of D and by D_v the completion of D at v (v is a place). U_v denotes the corresponding group of units in the ring of integers or the maximal compact subgroup of $GL(2,\mathbf{Q}_v)$ depending on whether D is ramified at v or not. Since D has class number one, we have

$$D^*_{\mathbf{A}} = D^* D^*_{\mathbf{R}} \prod_{v<\infty} U_v.$$

Suppose $U_3 = \cup\, a_i(1+P_3^2)$ where P_3 is the maximal ideal in the ring of integers in D_3.

Then $D^*_{\mathbf{A}} = \cup\, D^* g_i D^*_{\mathbf{R}} \prod_{v<\infty} U'_v$ where $U'_v = U_v$ if $v \neq 3$ and $U'_v = 1+P_3^2$

if $v = 3$ and g_i is the idele which has a_i in the third place and ones everywhere else.

In fact if we factor out units, then

$$D_{\mathbf{A}}^* = \bigcup_{i=1}^{6} D^* g_i D_{\mathbf{R}}^* \prod_{v<\infty} U'_v,$$

where $g_i = (1, 1, b_i, 1, \ldots)$ and the b_i's are coset representative such that $U_3 = \bigcup_{i=1}^{6} U(1+P_3^2)b_i$.

Denoted by N_{b_k}, the number of integral solutions of $N(a, b, c, d) = p$ such that

$$a+bj+\tfrac{1}{2}[c(1+i)]+\tfrac{1}{2}[d(1+i)j] \in b_k(1+P_3^2).$$

In $\mathbf{Q}_3$, -2 has a square root which lies in $1+(3)$. We denote this square root by $\sqrt{-2}$. Let $u = (1+j)/\sqrt{-2}$, $v = 1+iu$. Then the b_k's may be taken to be 1, u, v, v^2, uv, uv^2. Then from Jacquet-Langlands Theory, we get the following:

THEOREM 1. If $p \equiv -1 \pmod 3$, then $N_1 = N_v = N_{v^2} = 0$,

$$N_u = N_{uv} = N_{uv^2} = \tfrac{1}{3}(p+1).$$

If $p \equiv 1 \pmod 3$, then $N_u = N_{uv} = N_{uv^2} = 0$, $N_1 = \frac{1}{3}[p+1+2(2x-y)]$, $N_v = N_{v^2} = \frac{1}{3}[p+1-(2x-y)]$ where (x, y) is a solution of $x^2-xy+y^2 = p$ with $x \equiv 1 \pmod 3$, $y \equiv 0 \pmod 3$.

Remark. The fact that an integral solution of $N(a, b, c, d) = p$ contributes to N_{b_k} simply expresses some congruence condition on the solution. For instance a solution (a, b, c, d) contributes to N_1 simply means that $a+bj+\frac{1}{2}[c(1+i)]$ $+\frac{1}{2}[d(1+i)j]$ is in $1+P_3^2$ or 9 divides the norm of $a-1+bj+\frac{1}{2}[c(1+i)]$ $+\frac{1}{2}[d(1+i)j]$. This means that

$$9 \mid (a-1)^2+b^2+c^2+d^2+ac+bd-c.$$

Since $a^2+b^2+c^2+d^2+ac+bd = p$ we get the condition that $9 \mid p+1-2a-c$.

Proof of theorem. Let $E = Q(\sqrt{-3})$. E has class number 1 so

$$E_A^* = E^* \mathbf{C}^* \prod_{v<\infty} \theta_v^*$$

where θ_v is the ring of integers in E_v. In fact because of six units in E we have

$$E_A^* = E^* \mathbf{C}^* U' \text{ where } U' = \theta_2^*(1+3\theta_3) \prod_{v\neq 2,3} \theta_v^*$$

define a Grossencharacter x on E_A^* so that it is trivial on E^* and U' and maps $z \to (z\bar{z})^{-1/2}\bar{z}$ for $z \in \mathbf{C}$. Then the representation (see [3]) π_∞ of $GL(2, \mathbf{R})$ is just $\sigma(||^{1/2}, ||^{-1/2})$ (the irreducible subrepresentation obtained by inducing from the character

$$\begin{pmatrix} a_1 & x \\ 0 & a_2 \end{pmatrix} \to |a_1|^{1/2}|a_2|^{-1/2}.$$

The trivial representation of $D_{\mathbf{R}}^{*}$ is associated to $\sigma(||^{1/2}, ||^{-1/2})$ by the Weil representation.

The character χ_2 of $\mathbf{Q}_2$ is unramified. Moreover $\chi_2(2) = \chi_2(2(-½, -1, -½, -½,\ldots)) = -1$. Hence $\chi_2 = \eta_2 \cdot N_{E_2/\mathbf{Q}_2}$ where η_2 is an unramified character of $\mathbf{Q}_2^*$ and $N_{E_2/\mathbf{Q}_2}$ is the norm map from E_2 to $\mathbf{Q}_2$. Hence π_{χ_∞} is the representation of $GL_2(\mathbf{Q}_2)$ obtained by inducing from two unramified characters applied to the upper triangular group and so is of class 1 (see [1]). Hence there exists a vector e_2 in the space $\mathfrak{B}_2$ on which π_{χ_2} acts which is fixed by $GL_2(\mathbf{Z}_2)$. Moreover 2 remains a prime in E_2 so that $L_2(s,\chi)$ is of the form $1/[1-\chi_2(2)2^{-2s}]$. Let $a_2 = 0$. Similarly if $p \equiv -1(3)$, then p remains a prime in E_p. We let $a_p = 0$. χ_p is unramified, $\chi_p(p) = -1$ so $\chi_p = \eta_p . N_{E_p/\mathbf{Q}_p}$, η_p being unramified. Again there exists a vector e_p in the space $\mathfrak{B}_p$ on which π_{χ_p} acts, fixed by $GL_2(\mathbf{Z}_p)$. If $p \equiv 1(3)$, $p\theta_E = (\mathfrak{p}\bar{\mathfrak{p}})$ where $\mathfrak{p}$, $\bar{\mathfrak{p}}$ are two distinct conjugate ideals in the ring of integers in E. If $\mathfrak{p} = (x+\zeta y)$ where $\zeta = \frac{1}{2}(-1+\sqrt{-3})$, then $\bar{\mathfrak{p}} = (x+\bar{\zeta}y)$ and $p = x^2-xy+y^2$. Moreover if we choose (x,y) such that $x \equiv 1 \pmod 3$ and $y \equiv 0 \pmod 3$, then both $x+\zeta y$ and $x+\bar{\zeta}y$ when viewed as elements in E_3 lie in $1+3\theta_3$. In this case $\pi_{\chi_p} = \pi(\chi_{\mathfrak{p}}, \chi_{\bar{\mathfrak{p}}})$. $\chi_{\mathfrak{p}}$, $\chi_{\bar{\mathfrak{p}}}$ are unramified so the representation is class 1. Moreover

$$L_p(s,\chi) = 1/[1-\chi_{\mathfrak{p}}(\mathfrak{p})p^{-s}]1/[1-\chi_{\bar{\mathfrak{p}}}(\bar{\mathfrak{p}})p^{-s}]$$

$$= 1/[1-(\chi_{\mathfrak{p}}(\mathfrak{p})+\chi_{\bar{\mathfrak{p}}}(\bar{\mathfrak{p}}))p^{-s}+\chi_{\mathfrak{p}}(\mathfrak{p})\chi_{\bar{\mathfrak{p}}}(\bar{\mathfrak{p}})p^{-2s}]$$

$$\chi_{\mathfrak{p}}(\mathfrak{p})+\chi_{\bar{\mathfrak{p}}}(\bar{\mathfrak{p}}) = p^{-1/2}(2x-y).$$

Put $a_p = 2x-y$. Finally we have $\chi_3(\sqrt{-3}) = i$, $\chi_3(-\zeta) = -\zeta$ and χ_3 is trivial on $1+3\theta_3$. The second relationship tells us that χ_3 is not of the form $\eta_3 \cdot N_{E_3/\mathbf{Q}_3}$. Hence π_{χ_3} is supercuspidal. Let l be the smallest integer greater than zero such that χ_3 is trivial on $E_3' \cap (1+(\sqrt{-3})^{2l+1}\theta_3)$, E_3' being the norm one group. Here $l = 1$. Let π_3' be a representation of D_3^* such that the associated Weil representation of $GL_2(\mathbf{Q}_3)$ is π_{χ_3}. Then Proposition 3 of [5] tells us that dim $\pi_3' = (q+1)q^{l-1} = 4$ where q is the number of elements in the residue of $\mathbf{Q}_3$. Moreover, if χ_3' is the character of π_3' and α is that of π_{χ_3}, then $-\chi_3' =$ character of π_{χ_3} where this is defined. But by [5]

$$\alpha(1-\sqrt{-3}) = \chi_3(\delta)\ \mathrm{sgn}_{E_3/\mathbf{Q}_3}(\delta) \sum_{\substack{x \in U^0/U' \\ x \neq g_1, \bar{g}_1}} \chi_3(x)\ \mathrm{sgn}_{E_3/\mathbf{Q}_3}(\mathrm{tr}(g_1-x))$$

where $\mathrm{sgn}_{E_3/\mathbf{Q}_3}$ is the character of $\mathbf{Q}_3$ determined by the quadratic extension E_3, $\delta \in \mathbf{Q}_3$ is such that $(1-\sqrt{-3})/\delta$ has norm 1 and trace $(1-\sqrt{-3})/\delta \equiv 2 \pmod 3$, $g_1 = (1-\sqrt{-3})/\delta$. Here, $\mathbf{Q}_3(\sqrt{-3})$ has been imbedded in $GL_2(\mathbf{Q}_3)$ and the elements of $\mathbf{Q}_3(\sqrt{-3})$ identified with their images. Finally $U^i = E_3' \cap (1+\sqrt{-3})^{2i+1}\theta_3)$. Now

$$E_3' \cap ((1+\sqrt{-3})^3\theta_3) = E_3' \cap ((1+\sqrt{-3})^2\theta_3).$$

Moreover by Serre [4] we can see that

$$\frac{E_3' \cap (1+\sqrt{-3}\theta_3)}{E_3' \cap (1+\sqrt{-3})^2\theta_3} \cong \frac{1+\sqrt{-3}\theta_3}{1+3\theta_3}$$

and so has 3 elements, 1, g_1, $\bar{g}_1$. Hence

$$\alpha(1-\sqrt{-3}) = \chi_3(-2)\ \mathrm{sgn}_{E_3/\mathbf{Q}_3}(-2)\ \mathrm{sgn}_{E_3/\mathbf{Q}_3}\left(\mathrm{tr}\left(\frac{1-\sqrt{-3}}{-2}-1\right)\right) = -1. \quad (*)$$

The central character of π_3' is the same as that of π_{χ_3} which is $\chi_3 \cdot \mathrm{sgn}_{E_3/\mathbf{Q}_3}$. Hence $\pi_3'(-1) = \mathrm{I}$. Hence π_3' is a four-dimensional representation of $U_3/[\langle \pm 1\rangle(1+P_3^2)]$. This is a group of order 36 which has the following presentation: $\langle a,b,c;\ b^3 = 1,\ c^3 = 1,\ a^4 = 1,\ aba^{-1} = c,\ bc = cb,\ aca^{-1} = b^2\rangle$. It has 4 one-dimensional representations and two irreducible four-dimensional representations. If ρ_1 denotes the character of the subgroup generated by b and c which map $b \mapsto \omega$, $c \mapsto 1$ (ω a complex primitive cube root of 1) and ρ_2 the character which maps $b \mapsto \omega$, $c \mapsto \omega$, then the induced representations are inequivalent, irreducible four-dimensional representations. Denote their characters by χ and χ'. Then $\chi(b^2) = 1$, $\chi(bc) = -2$ and $\chi'(b^2) = -2$, $\chi'(bc) = 1$. Now (*) tells us that $\chi(b^2) = -1$. Hence $\chi_3' = \chi$. Here we have identified a, b, c with the cosets of u, $1+i$ and $1+ji$ respectively.

If V is the space on which π_3' acts, then

$$\pi' = \pi_{\chi_3} \otimes \pi_2' \otimes \bigotimes_{p>3} \pi_p \text{ acts on } \mathfrak{B}_2 \otimes V \otimes \bigotimes_{p>3} \mathfrak{B}_p$$

(restricted tensor product). However, by Jacquet-Langlands, π' occurs in the space $\mathcal{A}'$ of automorphic forms for $D_{\mathbf{A}}$ so there is an intertwining operator T: $\mathfrak{B}_2 \otimes V \otimes \bigotimes_{p>3} \mathfrak{B}_p \mapsto \mathcal{A}'$ which commutes with the $D_{\mathbf{A}}^*$ action. The image of the vectors $e_2 \otimes v \otimes \bigotimes_{p>3} e_p$, $v \in V$ lie in the space of functions on $D_{\mathbf{A}}^*$ which are left invariant by D^* and right invariant by $D_{\mathbf{R}}^* U_2(1+p_3^2) \prod_{p>3} U_p$. This space $\mathcal{F}$, we have seen in the beginning, is six-dimensional and contains a four-dimensional subspace on which the right regular action, restricted to D_3^*, is equivalent to π_3'. Let ϕ_x denote the characteristic function of the double coset

$$D^*(1,\ 1,\ x^{-1},\ 1,\ldots)D_{\mathbf{R}}^* \prod_{\nu<\infty} U_\nu',\ x \in U_3.$$

Then $\mathcal{F}$ is generated by $\langle \phi_1, \phi_u, \phi_v, \phi_{v^2}, \phi_{uv}, \phi_{uv^2}\rangle$, u, v being defined as before. Then the four-dimensional subspace in $\mathcal{A}'$ is generated by

$$\langle \phi_1 + \zeta\phi_v + \zeta^2\phi_{v^2},\ \phi_1 + \zeta^2\phi_v + \zeta\phi_{v^2},\ \phi_u + \zeta\phi_{uv} + \zeta^2\phi_{uv^2},$$
$$\phi_u + \zeta^2\phi_{uv} + \zeta\phi_{uv^2}\rangle.$$

If p is an odd prime, then the corresponding Hecke operator commutes with the right regular action of U_3 on this space. Since this action is irreducible the Hecke operators act as scalars. The given functions are, therefore, eigenfunctions for the

T_p with eigenvalues a_p. By Jacquet-Langlands $a_p = a'_p$. We need to compute a'_p. We have

$$a'_p(\phi_1+\zeta\phi_v+\zeta^2\phi_{v^2}) = T_p(\phi_1+\zeta\phi_v+\zeta^2\phi_{v^2}).$$

Evaluating at 1 we get

$$\begin{aligned} a'_p &= [T_p(\phi_1+\zeta\phi_v+\zeta^2\phi_{v^2})](1) \\ &= \sum_{x\in \mathbf{Z}/p\mathbf{Z}} \phi_1\left[\begin{pmatrix} p & x \\ 0 & 1\end{pmatrix}\right]+\phi_1\left[\begin{pmatrix} 1 & 0 \\ 0 & p\end{pmatrix}\right] \\ &\quad +\zeta \sum_{x\in \mathbf{Z}/p\mathbf{Z}} \phi_v\left[\begin{pmatrix} p & x \\ 0 & 1\end{pmatrix}\right] \\ &\quad +\zeta\phi_v\left[\begin{pmatrix} 1 & 0 \\ 0 & p\end{pmatrix}\right]+\zeta^2 \sum_{x\in \mathbf{Z}/p\mathbf{Z}} \phi_{v^2}\left[\begin{pmatrix} p & x \\ 0 & 1\end{pmatrix}\right]+\zeta^2\phi_{v^2}\left[\begin{pmatrix} 1 & 0 \\ 0 & p\end{pmatrix}\right] \\ &= N_1+\zeta^2N_v+\zeta N_{v^2};\ \text{see [6]}. \end{aligned}$$

Similarly evaluating at v^2 we get $a'_p = N_1+\zeta N_v+\zeta^2 N_{v^2}$.
Evaluating at u, uv we get $0 = N_u+\zeta N_{uv}+\zeta^2 N_{uv^2} = N_u+\zeta^2 N_{uv}+\zeta N_{uv^2}$.
Suppose now $p \equiv -1(3)$. Then $a_p = a'_p = 0$ so we get

$$N_1 = N_v = N_{v^2},\ N_u = N_{uv} = N_{uv^2}.$$

Moreover, we claim that $N_1 = 0$. For if the solution (a, b, c, d) contributes to N_1, then $9/p+1-2a-c$. Hence $3|a-c$. But

$$\begin{aligned} p = a^2+b^2+c^2+d^2+ac+bd &= (a-c)^2+(b-d)^2+3ac+3bd \\ &\equiv (b-d)^2 \pmod 3. \end{aligned}$$

Therefore $(p/3) = 1$ which is a contradiction so $N_1 = 0$. Since the sum of the N's is $p+1$, we get

$$N_u = \tfrac{1}{3}(p+1).$$

On the other hand suppose $p \equiv 1 \pmod 3$, x, y being as before we get

$$\begin{aligned} 2x-y &= N_1+\zeta^2N_v+\zeta N_{v^2} = N_1+\zeta N_v+\zeta^2 N_{uv^2} \\ 0 &= N_u+\zeta^2 N_{uv}+\zeta N_{uv^2} = N_u+\zeta N_{uv}+\zeta^2 N_{uv^2}, \end{aligned}$$

so $N_u = N_{uv} = N_{uv^2}$, $N_v = N_{v^2}$ and $2x-y = N_1-N_v$. We claim that N_u is zero. For if (a, b, c, d) is a solution contributing towards N_u, then 9 divides $p+2-2a-2b-c-d$. Hence 3 divides $a+b-c-d$, i.e., $(a-c) = d-b \pmod 3$. But $p = (a-c)^2+(d-b)^2+3ac+3bd \equiv 2(a-c)^2 \pmod 3$ which is a contradiction. Hence $N_1+2N_v = p+1$ and $N_1-N_v = 2x-y$. This gives the required solution and completes the proof of the theorem.

The same argument may be repeated for definite quaternion algebras with discriminant 7, the essential conditions involved, namely that the class numbers of $\mathbf{Q}(\sqrt{-7})$ and the quaternion algebra are one, being the same. Such a quaternion algebra is generated over $\mathbf{Q}$ by $1, i, j, ij$ such that $i^2 = -7$ and $j^2 = -1$. The unique

maximal order is $Z+Zj+\frac{1}{2}[Z(1+i)]+\frac{1}{2}[Z(1+i)j]$ and the associated norm form $N(a, b, c, d) = a^2+b^2+2c^2+2d^2+ac+bd$. If U be the units in the maximal order and "a", a 48th root of unity in $Q_7(\sqrt{-7}) = Q_7(i)$ then $U_7 = \bigcup_{i=0}^{11} U\alpha^i(1+P_7)$. We denote as before by N_i the number of integral solutions of $N(a, b, c, d) = p$ such that $a+bj+\frac{1}{2}[c(1+i)]+\frac{1}{2}[d(1+i)j]$ is in $a^i(1+P_7)$. Making a closer use of the character tables in [5] we get

THEOREM 2. If $(p/7) = -1$, then $N_i = 0$ if i is even and $N_i = \frac{1}{2}(p+1)$ if i is odd.

If $(p/7) = 1$, then $N_i = 0$ if i is odd, $N_i = \frac{1}{2}[p+1+(2x+y)]$ if 4 divides i and $N_i = \frac{1}{2}[p+1-(2x+y)$ if i is even but not divisible by 4; here (x,y) is any solution of $x^2+xy+2y^2 = p$ such that $(2x+y)/p = 1$.

The notation $(/p)$ is used for the Legendre symbol.

References

[1] Gelbart S, Automorphic forms on adele groups, *Ann. Math. Stud.* No. 83 1975 (Princeton: University Press)

[2] Hecke E, Werke, *Vandenhoeck and Ruprecht*, Gottingen 1959

[3] Jacqet H and Langlands R P, Automorphic forms on $GL(2)$, *Lecture notes in mathematics* 114 (Berlin: Springer Verlag)

[4] Serre J P, *Corps Locaux* (Paris: Hermann)

[5] Shimizu H, Some examples of new forms, *J. Fac. Sci. Univ. Tokyo Sec. IA* **24** 97–113

[6] Tandon R, The Hecke theory of $GL(2)$ and quadratic forms, *J. Indian Math. Soc.* **40** (1976) 87–122

Proc. Indian Acad. Sci. (Math. Sci.), Vol. 95, No. 2, December 1986, pp. 133–140.

On a class of bilateral generating functions for certain special functions

A K CHONGDAR
Department of Mathematics, Bangabasi Evening College, Calcutta 700 009, India

MS received 3 October 1986

Abstract. A general theorem unifying a novel class of bilateral generating functions of certain special functions is established. A number of applications of the theorem are also given.

Keywords. Bilateral generating functions; special functions.

1. Introduction

Several attempts have been made by many researchers [2, 4, 9, 11, 17] to formulate theories in connection with the unification of different classes of bilateral generating functions of certain special functions found in the literature. The present note is a further attempt to establish a general theorem on the unification of a novel class of bilateral generating functions of various special functions e.g. Laguerre, modified Laguerre, generalized Bessel, Gegenbauer and modified Jacobi polynomials. In fact we have obtained the following theorem:

THEOREM 1. For a set of functions $\{S_n^k(x) \mid n = 0, 1, 2, \ldots\}$ generated by

$$\sum_{n=0}^{\infty} A_n(m,k)\, S_{n+m}^{k-n}(x) t^n = \frac{\{f(x,t)\}^k g(x,t)}{\{h(x,t)\}^m}\, S_m^k\{q(x,t)\}, \tag{1}$$

where m is a non-negative integer, A_n the arbitrary constant and f, g, h, q are arbitrary functions of x and t, let

$$F(x,t) = \sum_{n=0}^{\infty} a_n\, S_{n+r}^{k-n}(x) t^n, \tag{2}$$

then the following bilateral generating relation for $S_r^k(x)$ holds:

$$\sum_{n=0}^{\infty} S_{n+r}^{k-n}(x)\sigma_n(y) t^n = \frac{\{f(x,t)\}^k g(x,t)}{\{h(x,t)\}^r}\, F\left(q(x,t), \frac{yt}{f(x,t)h(x,t)}\right), \tag{3}$$

where

$$\sigma_n(y) = \sum_{p=0}^{n} a_p\, A_{n-p}(r+p, k-p) y^p.$$

During the application of the above theorem, a large number of bilateral generating functions for various special functions are pointed out.

Proof of theorem 1

We have

$$\sum_{n=0}^{\infty} S_{n+r}^{k-n}(x)\sigma_n(y)t^n = \sum_{p=0}^{\infty} a_p(yt)^p \sum_{n=0}^{\infty} A_n S_{n+p+r}^{k-n-p}(x)t^n$$

$$= \frac{\{f(x,t)\}^k g(x,t)}{\{h(x,t)\}^r} \sum_{p=0}^{\infty} a_p \left(\frac{yt}{f(x,t)\,h(x,t)}\right)^p S_{r+p}^{k-p}(q(x,t))$$

$$= \frac{\{f(x,t)\}^k g(x,t)}{\{h(x,t)\}^r} F\left(q(x,t), \frac{yt}{f(x,t)\,h(x,t)}\right).$$

This completes the proof of the theorem.

COROLLARY 1. If we put $r = 0$, we get the following result. If

$$F(x,t) = \sum_{n=0}^{\infty} a_n S_n^{k-n}(x)t^n,$$

then

$$\sum_{n=0}^{\infty} S_n^{k-n}(x)\sigma_n(y)t^n = \{f(x,t)\}^k g(x,t)\, F\left(q(x,t), \frac{yt}{f(x,t)\,h(x,t)}\right).$$

where

$$\sigma_n(y) = \sum_{p=0}^{n} a_p A_{n-p} y^p.$$

In the next section we proceed to give the various applications of the theorem in the field of certain special functions.

2. Applications

(i) *On Laguerre polynomials*

We first consider the Laguerre polynomials satisfying the following generating relation [1, 7]

$$(1+t)^k \exp(-xt)\, L_r^{(k)}(x(1+t)) = \sum_{n=0}^{\infty} \frac{(r+1)_n}{n!} L_{n+r}^{(k-n)}(x)t^n. \tag{4}$$

The relation (4) is of type (1) with,

$$f(x,t) = 1+t,\ g(x,t) = \exp(-xt),\ h(x,t) = 1,\ q(x,t) = x(1+t)$$

and $A_n = \dfrac{(r+1)_n}{n!}$.

Now using the theorem, we get the following result on bilateral generating relation involving Laguerre polynomials:

THEOREM 2. If there exists a generating relation of the form

$$F(x,t) = \sum_{n=0}^{\infty} a_n L_{n+r}^{(k-n)}(x)t^n, \tag{5}$$

then

$$\sum_{n=0}^{\infty} L_{n+r}^{(k-n)}(x)\sigma_n(y)t^n = (1+t)^k \exp(-xt)\, F\left(x(1+t), \frac{yt}{1+t}\right), \tag{6}$$

where

$$\sigma_n(y) = \sum_{p=0}^{n} a_p \frac{(r+p+1)_{n-p}}{(n-p)!}\, y^p.$$

COROLLARY 2. If we put $r = 0$ in the above theorem we get the following result. If

$$F(x,t) = \sum_{n=0}^{\infty} a_n L_n^{(k-n)}(x)t^n,$$

then

$$\sum_{n=0}^{\infty} L_n^{(k-n)}(x)\sigma_n(y)t^n = (1+t)^k \exp(-xt)\, F\left(x(1+t), \frac{yt}{1+t}\right),$$

where

$$\sigma_n(y) = \sum_{p=0}^{n} a_p \frac{(p+1)_{n-p}}{(n-p)!}\, y^p.$$

(ii) *On modified Laguerre polynomials*

We now consider the following well-known generating function for the modified Laguerre polynomials [10]

$$\sum_{n=0}^{\infty} \frac{(r+1)_n}{n!} f_{n+r}^{k-n}(x)t^n = (1+t)^{k-1} \exp\left(\frac{xt}{1+t}\right) f_r^k\left(\frac{x}{1+t}\right). \tag{7}$$

The relation (7) is of type (1) with

$$f(x,t) = 1+t,\ g(x,t) = (1+t)^{-1} \exp\left(\frac{xt}{1+t}\right),$$

$$h(x,t) = 1,\ q(x,t) = \frac{x}{1+t} \quad \text{and } A_n = \frac{(r+1)_n}{n!}.$$

Therefore by the application of our theorem we get the following result on bilateral generating relation involving modified Laguerre polynomials.

THEOREM 3. If there exists a generating relation of the form

$$F(x,t) = \sum_{n=0}^{\infty} a_n f_{n+r}^{k-n}(x) t^n, \tag{8}$$

then

$$\sum_{n=0}^{\infty} f_{n+r}^{k-n}(x)\, \sigma_n(y) t^n = (1+t)^{k-1} \exp\left(\frac{xt}{1+t}\right) F\left(\frac{x}{1+t}, \frac{yt}{1+t}\right)$$

where

$$\sigma_n(y) = \sum_{p=0}^{n} a_p \frac{(r+p+1)_{n-p}}{(n-p)!} y^p. \tag{9}$$

COROLLARY 4. Putting $r = 0$ in the above theorem, we get the known result derived by Ghosh [15].

(iii) *On Bessel polynomials*

We now consider the generalized Bessel polynomials satisfying the following generating function [8]

$$\sum_{n=0}^{\infty} \frac{\beta^n}{n!} Y_{n+r}^{(k-n)}(x) t^n = (1-xt)^{1-k-r} \exp(\beta t)\, Y_r^{(k)}\left(\frac{x}{1-xt}\right). \tag{10}$$

The relation (10) is of the same type (1) with

$$f(x,t) = (1-xt)^{-1},\ g(x,t) = (1-xt)\exp(\beta t),\ h(x,t) = 1-xt,$$

$$q(x,t) = \frac{x}{1-xt}, \quad \text{and } A_n = \beta^n/n!.$$

Therefore by using our theorem, we get the following result involving Bessel polynomials.

THEOREM 4. If there exists a generating relation of the form

$$F(x,t) = \sum_{n=0}^{\infty} a_n Y_{n+r}^{(k-n)}(x) t^n \tag{11}$$

then

$$\sum_{n=0}^{\infty} Y_{n+r}^{(k-n)}(x)\sigma_n(y)t^n = (1-xt)^{1-k-r}\exp(\beta t)F\left(\frac{x}{1-xt}, yt\right)$$

where

$$\sigma_n^{\hat{}}(y) = \sum_{p=0}^{n} a_p \frac{\beta^{n-p}}{(n-p)!}y^p. \tag{12}$$

COROLLARY 4. Putting $r = 0$ in the above theorem we get the interesting result found derived with some misprints in [6].

(iv) *On Gegenbauer polynomials*

Next we consider the Gegenbauer polynomials satisfying the following generating relation [5]

$$\sum_{n=0}^{\infty} \binom{n+r}{r} \frac{(1-r-2k)_n}{(1-k)_n} C_{n+r}^{k-n}(x)t^n$$

$$= [1+4tx+4t^2(x^2-1)]^{k-1/2}\, C_r^k(x+2t(x^2-1)). \tag{13}$$

The relation (13) is of type (1) with

$$f(x,t) = 1+4tx+4t^2(x^2-1),\ g(x,t) = [1+4tx+4t^2(x^2-1)]^{-1/2}$$

$$h(x,t) = 1,\ q(x,t) = x+2t(x^2-1),\ A_n = \binom{n+r}{r}\frac{(1-r-2k)_n}{(1-k)_n}.$$

Now using theorem 1, we get the following result on the bilateral generating function of Gegenbauer polynomials

THEOREM 5. If there exists a generating relation of the form

$$F(x,t) = \sum_{n=0}^{\infty} a_n C_{n+r}^{k-n}(x)t^n \tag{14}$$

then

$$\sum_{n=0}^{\infty} C_{n+r}^{k-n}(x)\sigma_n(y)t^n = [1+4tx+4t^2(x^2-1)]^{k-1/2} \times$$

$$F\left(x+2t(x^2-1), \frac{yt}{1+4tx+4t^2(x^2-1)}\right) \tag{15}$$

where

$$\sigma_n(y) = \sum_{p=0}^{n} a_p \binom{n+r}{r+p} \frac{(1-r+p-2k)_{n-p}}{(1-k+p)_{n-p}} y^p.$$

COROLLARY 5. Putting $r = 0$ in theorem 5, we get the result derived in [14].

(v) *On Jacobi polynomials*

Next we consider the Jacobi polynomials satisfying the following generating relation [3]

$$\sum_{n=0}^{\infty} \frac{(r+1)_n}{n!} P_{n+r}^{(\alpha,k-n)}(x)t^n \tag{16}$$

$$= (1-t)^k \left\{1-(1+x)\frac{t}{2}\right\}^{-1-\alpha-k-r} P_r^{(\alpha,k)}\left(\frac{x-\frac{t}{2}(1+x)}{1-\frac{t}{2}(1+x)}\right).$$

The relation (16) is of type (1) with

$$f(x,t) = \frac{1-t}{1-\frac{t}{2}(1+x)}, \quad g(x,t) = \left\{1-\frac{t}{2}(1+x)\right\}^{-1-\alpha},$$

$$h(x,t) = \left\{1-\frac{t}{2}(1+x)\right\},$$

$$q(x,t) = \left\{1-\frac{t}{2}(1+x)\right\}\Big/\left\{1-\frac{t}{2}(1+x)\right\}, \quad A_n = (r+1)_n/n!.$$

Now using theorem 1, we get the following result on the bilateral generating relation involving Jacobi polynomials.

THEOREM 6. If there exists a generating relation of the form

$$F(x,t) = \sum_{n=0}^{\infty} a_n P_{n+r}^{(\alpha,k-n)}(x)t^n, \tag{17}$$

then

$$\sum_{n=0}^{\infty} P_{n+r}^{(\alpha,k-n)}(x)\sigma_n(y)t^n \tag{18}$$

$$= (1-t)^k \left\{1-\frac{t}{2}(1+x)\right\}^{-1-\alpha-k-r} F\left(\frac{x-\frac{t}{2}(1+x)}{1-\frac{t}{2}(1+x)}, \frac{yt}{1-t}\right),$$

where

$$\sigma_n(y) = \sum_{p=0}^{n} a_p \frac{(r+p+1)_{n-p}}{(n-p)!} y^p.$$

COROLLARY 6. Putting $r = 0$ in the above theorem we get the known result derived in [12].

If in place of relation (16) we consider the following [16]

$$\sum_{n=0}^{\infty} \frac{(r+1)_n}{n!} P_{n+r}^{(k-n,\beta)}(x)t^n = (1+t)^k \left\{1+\frac{t}{2}(1-x)\right\}^{-1-k-\beta-r} \times$$
$$P_r^{(k,\beta)}\left(\frac{x-\frac{t}{2}(1-x)}{1+\frac{t}{2}(1-x)}\right). \tag{19}$$

we get the following result.

THEOREM 7. If there exists a generating relation of the form

$$F(x,t) = \sum_{n=0}^{\infty} a_n P_{n+r}^{(k-n,\beta)}(x)t^n, \tag{20}$$

then

$$\sum_{n=0}^{\infty} P_{n+r}^{(k-n,\beta)}(x)\sigma_n(y)t^n = (1+t)^k \left\{1+\frac{t}{2}(1-x)\right\}^{-1-k-\beta-r} \times$$
$$F\left(\frac{x-\frac{t}{2}(1-x)}{1+\frac{t}{2}(1-x)}, \frac{yt}{1+t}\right),$$

where

$$\sigma_n(y) = \sum_{p=0}^{n} a_p \frac{(r+p+1)_{n-p}}{(n-p)!} y^p.$$

COROLLARY 7. Putting $r = 0$ in the above theorem we get the result which has been derived in [13].

Conclusion

From the above discussion it is clear that one may apply theorem 1 in the case of other polynomials and functions existing in the field of special functions to obtain the bilateral generating relation involving the special function under consideration.

References

[1] Al-Salam W A, Operational representations for the Laguerre and other polynomials, *Duke Math. J.* **31** (1964) 127–142

[2] Basu D K, On the unification of a class of bilateral generating relations for special functions, *Math. Balk.* **6** (1976) 21–29

[3] Chakraborty A B, Group-theoretic study of certain generating functions of Jacobi polynomials, *J. Indian Inst. Sci.* **64B** (1983) 97–103

[4] Chatterjea S K, Unification of a class of bilateral generating relations for certain special functions, *Bull. Calif. Math. Soc.* **67** (1975) 115–127

[5] Chongdar A K, Group-theoretic study of certain generating functions of Gegenbauer polynomials, *Orissa Math. J.* **2** (1983) 71–85

[6] Chongdar A K, Some generating functions of modified Bessel polynomials from the viewpoint of Lie-group, *Int. J. Math. Math. Sci.* **7** (1984a) 823–825

[7] Chongdar A K, Some generating functions, involving Laguerre polynomials, *Bull. Calif. Math. Soc.* **76** (1984b) 262–269

[8] Chongdar A K, Some generating functions involving generalized Bessel polynomials, *Bull. Calif. Math. Soc.* **77** (1985) 330–339

[9] Chongdar A K, A general theorem on the unification of a class of bilateral generating functions for certain special functions, *Bull. Calif. Math. Soc.* (communicated) (1986a)

[10] Chongdar A K, A Lie algebra and generating functions of modified Laguerre polynomials, *Bull. Calif. Math. Soc.* **78** (1986b) 219–226

[11] Chongdar A K, on the unification of a class of bilateral generating functions for certain special functions, *Tamkang J. Math.* **18** (1987)

[12] Chongdar A K and Chatterjea S K, On a class of trilateral generating relations with Tchebycheff polynomials from the viewpoint of one parameter group of continuous transformations, *Bull. Calif. Math. Soc.* **73** (1981) 127–140

[13] Das S, On partial differential operators for Jacobi polynomials, *Pure Math. Manuscript* **3** (1984) 95–104

[14] Das S and Chatterjea S K, On a partial differential operator for Gegenbauer polynomials, *Bull. Calif. Math. Soc.* **76** (1984) 351–361

[15] Ghosh B, Some theorems on generating functions of modified Laguerre polynomials, *Bull. Calif. Math. Soc.* **79** (1987) 41–47

[16] Ghosh B, Group-theoretic origins of certain generating functions of Jacobi polynomials II, to appear in *Pure Math. Manuscript* **5** (1986).

[17] Singhal J P and Srivastava H M, A class of bilateral generating functions for certain classical polynomials, *Pacific J. Math.* **42** (1972) 355–362

Proc. Indian Acad. Sci. (Math. Sci.), Vol. 95, No. 2, December 1986, pp. 141–153.

Asymptotic properties of solutions of difference equations

J POPENDA

Institute of Mathematics, Technical University, 60–965 Poznań, Poland

MS received 18 November 1986

Abstract. Sufficient conditions for some m-th order finite difference equations are presented which have a solution behaving in a precisely specified way like a given polynomial.

Keywords. Difference equation; asymptotic behaviour.

1. Introduction

In this paper we examine the asymptotic behaviour of solutions of certain classes of m-th order difference equations. Motivated with the work of Trench [9] on differential equations we give conditions which imply that the equation

$$\Delta^m y_n + f(n, y_n) = 0 \quad m \geq 1, n \in N \tag{E}$$

has a solution which behaves like a given polynomial of degree $< m$ as $n \to \infty$.

Some asymptotic properties of solutions of second order difference equations were considered in [3–5, 8]. Results similar to those contained in this paper were presented in [1], [2], [6], [7]. Here $y_n = y(n), R\colon = (-\infty, \infty)$, $R_0\colon = [0, \infty)$, $R_+\colon = (0, \infty)$, $N\colon = \{n_0, n_0+1, \ldots\}$, n_0 is a given non-negative integer. For a function $x\colon N \to R$, we define the difference operators Δ^i as follows

$$\Delta^0 x_n = x_n, \; \Delta^k x_n = \Delta(\Delta^{k-1} x_n) = \Delta^{k-1} x_{n+1} - \Delta^{k-1} x_n, \; k \geq 1.$$

We write $x_n = O(z_n)$ and $x_n = o(z_n)$ to indicate that

$$\varlimsup_{n \to \infty} \left| \frac{x_n}{z_n} \right| < \infty \text{ or } \lim_{n \to \infty} \frac{x_n}{z_n} = 0$$

respectively. By a solution of (E) we mean any function x defined on N, which fulfils (E) for all sufficiently large n. Note that the above definition of the solution is different from this where x fulfils (E) for all $n \in N$.

2. Main result

THEOREM

Let $q: N \to R_+$ be nonincreasing, p be a given polynomial of degree $< m$ and suppose that there is a constant M such that the functions $f(n,.)$, $n \in N$ are continuous on the sets

$$U_n := \{u: |u-p_n| \leq Mq_n\}$$

and

$$|f(n,u)-f(n,v)| \leq g_n|u-v|, \tag{1}$$

if $u,v \in U_n$, where $g: N \to R_0$ and

$$\sum_{j=n_0}^{\infty} j^{m-1} g_j q_j \tag{2}$$

is convergent,

$$\overline{\lim_{n\to\infty}} \frac{1}{q_n} \sum_{j=n}^{\infty} j^{m-1} g_j q_j = c_1 < (m-1)!. \tag{3}$$

Suppose also that

$$\sum_{j=n_0}^{\infty} j^{m-1} f(j,p_j) \tag{4}$$

converges—perhaps conditionally—and

$$\overline{\lim_{n\to\infty}} \frac{1}{q_n} \left| \sum_{j=n}^{\infty} f(j,p_j) \right| = c_2, \tag{5}$$

where

$$c_2 + c_1 M < M(m-1)!. \tag{6}$$

Then (E) has a solution x which satisfies the asymptotic relation

$$\Delta^i x_n = \Delta^i p_n + O(n^{-i} q_n), \; O \leq i \leq m-1. \tag{AR}$$

Proof. Let $m(q)$ be the Banach space of sequences $h: N \to R$ such that $h_n = O(q_n)$ with the norm

$$||h|| = \sup_{n \geq n_0} \{ |h_n|/q_n \}. \tag{7}$$

Let

$$m_M(q): = \{h \in m(q): ||h|| \leq M\}.$$

Take $\varepsilon > 0$ such that

$$c_2 + c_1 M + \varepsilon < (m-1)!M, \; c_1 + \frac{\varepsilon}{2M} < (m-1)!. \tag{8}$$

By (3) and (5) there exist $n_1, n_2 \in N$ such that

$$\frac{1}{q_n} \sum_{j=n}^{\infty} j^{m-1} g_j q_j \le c_1 + \frac{\varepsilon}{2M} = c_3 \tag{9}$$

for all $n \ge n_1$, and

$$\frac{1}{q_n} \left| \sum_{j=n}^{\infty} f(j, p_j) \right| \le c_2 + \frac{\varepsilon}{2} = c_4 \tag{10}$$

for all $n \ge n_2$. Take $n_3 = \max\{n_1, n_2, m-1\}$ and define operator T, $\hat{h} = Th$ by

$$\hat{h}_n = \begin{cases} 0 & \text{for } n < n_3 \\ (-1)^{m-1} \displaystyle\sum_{j=n}^{\infty} \frac{(j+m-1-n)^{(m-1)}}{(m-1)!} f(j, p_j + h_j) & \text{for } n \ge n_3 \end{cases} \tag{11}$$

for $h \in m_M(q)$, where by $(n)^{(k)}$ we indicate product $n(n-1) \ldots (n-k+1)$ for $k \ge 1$, $(n)^{(0)} = 1$.

We will show that T is a contraction mapping of $m_M(q)$ into itself. Let $h \in m_M(q)$. We must first prove that the series in (11) converges. Denote

$$I(n, h) = \sum_{j=n}^{\infty} j^{m-1} f(j, p_j + h_j). \tag{12}$$

We note that assumption (4) implies that

$$I(n, O) = \sum_{j=n}^{\infty} j^{m-1} f(j, p_j) \tag{13}$$

converges. By virtue of (1), (2) and (7) we get for any $u, v \in m_M(q)$

$$\sum_{j=n}^{\infty} |j^{m-1}[f(j, p_j + u_j) - f(j, p_j + v_j)]| \le \sum_{j=n}^{\infty} j^{m-1} g_j |u_j - v_j|$$

$$\le \|u - v\| \sum_{j=n}^{\infty} j^{m-1} g_j q_j \le c_5 - \text{constant.}$$

Therefore the series

$$\sum_{j=n}^{\infty} j^{m-1}[f(j, p_j + u_j) - f(j, p_j + v_j)] \tag{15}$$

converges for any u, $v \in m_M(q)$. Taking $u = h$, $v \equiv 0$, by convergence of the series

(13) and (15) we obtain convergence of (12) for arbitrary $h \in m_M(q)$. It should be observed that the sequence

$$\left\{\frac{(j+m-1-n)^{(m-1)}}{j^{m-1}}\right\}_{j=n}^{\infty} = \left\{\prod_{k=0}^{m-2}\left(1+\frac{(m-1-n-k)}{j}\right)\right\}_{j=n}^{\infty}$$

is bounded and increasing.

Furthermore

$$\lim_{j\to\infty}\prod_{k=0}^{m-2}\left(1+\frac{(m-1-n-k)}{j}\right) = 1. \tag{16}$$

Using this fact and the convergence of (12), by Abel's test we obtain the convergence of the series

$$\sum_{j=n}^{\infty}\frac{(j+m-1-n)^{(m-1)}}{j^{m-1}}\, j^{m-1} f(j, p_j+h_j)$$

$$= \sum_{j=n}^{\infty}(j+m-1-n)^{(m-1)} f(j, p_j+h_j). \tag{17}$$

We shall estimate $\hat{h}_n$ for $n \ge n_3$. Differentiating (12) we obtain

$$\Delta I(n,h) = -n^{m-1} f(n, p_n+h_n). \tag{18}$$

On account of (18), summing by parts we have

$$\sum_{j=n}^{t}(j+m-1-n)^{(m-1)} f(j, p_j+h_j)$$

$$= -\sum_{j=n}^{t}(j+m-1-n)^{(m-1)}\frac{\Delta I(j,h)}{j^{m-1}}$$

$$= -\sum_{j=n}^{t}\left[\prod_{k=0}^{m-2}\left(1+\frac{(m-1-n-k)}{j}\right)\right]\Delta I(j,h)$$

$$= -\left[\prod_{k=0}^{m-2}\left(1+\frac{(m-1-n-k)}{t}\right)\right] I(t+1,h)$$

$$+\left[\prod_{k=0}^{m-2}\left(1+\frac{(m-1-n-k)}{n}\right)\right] I(n,h)$$

$$+\sum_{j=n}^{t-1} I(j+1,h)\,\Delta\prod_{k=0}^{m-2}\left(1+\frac{(m-1-n-k)}{j}\right).$$

Passing with t to infinity, by (16), convergence of the series (17), and

$$\lim_{t\to\infty} I(t+1,h)=0$$

we get convergence of the series

$$\sum_{j=n}^{\infty} I(j+1,h)\Delta\prod_{k=0}^{m-2}\left(1+\frac{(m-1-n-k)}{j}\right)$$

and equality

$$=\sum_{j=n}^{\infty}(j+m-1-n)^{(m-1)}f(j,p_j+h_j)$$

$$=\frac{(m-1)!I(n,h)}{n^{m-1}}$$

$$+\sum_{j=n}^{\infty} I(j+1,h)\Delta\prod_{k=0}^{m-2}\left(1+\frac{m-1-n-k}{j}\right). \tag{19}$$

Hence we see that to obtain an estimation of (17) we need any bound of $I(n,h)$. To this we denote

$$s_n=\sup_{k\geq n}\frac{1}{q_k}\left[M\sum_{j=k}^{\infty}j^{m-1}g_jq_j+\left|\sum_{j=k}^{\infty}j^{m-1}f(j,p_j)\right|\right]. \tag{20}$$

Therefore by (14) and (20) we get

$$|I(n,h)|=\left|\sum_{j=n}^{\infty}j^{m-1}[f(j,p_j+h_j)-f(j,p_j)]+\sum_{j=n}^{\infty}j^{m-1}f(j,p_j)\right|$$

$$\leq\|h\|\sum_{j=n}^{\infty}j^{m-1}g_jq_j+\left|\sum_{j=n}^{\infty}j^{m-1}f(j,p_j)\right|\leq q_n\left\{\frac{1}{q_n}\left[M\sum_{j=n}^{\infty}j^{m-1}g_jq_j\right.\right.$$

$$\left.\left.+\left|\sum_{j=n}^{\infty}j^{m-1}f(j,p_j)\right|\right]\right\}.$$

$$\leq q_ns_n \text{ for } n\geq n_0. \tag{21}$$

Applying (8), (9), and (10) to (20) yields

$$s_n=M\sup_{k\geq n}\frac{1}{q_k}\sum_{j=k}^{\infty}j^{m-1}g_jq_j+\sup_{k\geq n}\frac{1}{q_k}\left|\sum_{j=k}^{\infty}j^{m-1}f(j,p_j)\right|$$

$$\leq M\left(c_1+\frac{\varepsilon}{2M}\right)+c_2+\frac{\varepsilon}{2}<(m-1)!M. \tag{22}$$

So by (19), (21), (22) we can estimate $|\hat{h}_n|$ as follows

$$|\hat{h}_n| = \left| \sum_{j=n}^{\infty} \frac{(j+m-1-n)^{(m-1)}}{(m-1)!} f(j, p_j + h_j) \right|$$

$$= \frac{1}{(m-1)!} \left| \frac{(m-1)!\, I(n,h)}{n^{m-1}} \right.$$

$$\left. + \sum_{j=n}^{\infty} I(j+1,h) \Delta \prod_{k=0}^{m-2} \left(1 + \frac{(m-1-n-k)}{j}\right) \right|$$

$$\le \frac{1}{(m-1)!} \left[\frac{(m-1)!\, q_n s_n}{n^{m-1}} + \sum_{j=n}^{\infty} q_{j+1} s_{j+1} \Delta \prod_{k=0}^{m-2} \left(1 + \frac{m-1-n-k}{j}\right) \right]$$

$$\le \frac{1}{(m-1)!} \left[\frac{(m-1)!\, q_n s_n}{n^{m-1}} + q_n s_n \sum_{j=n}^{\infty} \Delta \prod_{k=0}^{m-2} \left(1 + \frac{m-1-n-k}{j}\right) \right].$$

Since

$$\sum_{j=n}^{\infty} \Delta \prod_{k=0}^{m-2} \left(1 + \frac{m-1-n-k}{j}\right) = \lim_{t \to \infty} \left[\prod_{k=0}^{m-2} \left(1 + \frac{m-1-n-k}{t+1}\right) \right.$$

$$\left. - \prod_{k=0}^{m-2} \left(1 + \frac{m-1-n-k}{n}\right) \right]$$

$$= 1 - \frac{(m-1)!}{n^{m-1}}$$

then

$$|\hat{h}_n| \le \frac{1}{(m-1)!} [s_n q_n] < q_n M, \quad n \ge n_0. \tag{23}$$

Hence we infer $\hat{h} \in m_M(q)$. This means that the operator T defined by (11) maps $m_M(q)$ into itself. We shall show that T is a contraction mapping. Assume $u, v \in m_M(q)$ and $\hat{u} = Tu$, $\hat{v} = Tv$. Then by (1), (7) taking into account that $j \ge n \ge n_3 \ge m-1$ and $(j)^{(k)} \le j^k$ we deduce

$$|\hat{u}_n - \hat{v}_n| = \frac{1}{(m-1)!} \left| \sum_{j=n}^{\infty} (j+m-1-n)^{(m-1)} f(j, p_j + u_j) \right.$$

$$\left. - \sum_{j=n}^{\infty} (j+m-1-n)^{(m-1)} f(j, p_j + v_j) \right|$$

$$\le \frac{1}{(m-1)!} \sum_{j=n}^{\infty} (j+m-1-n)^{(m-1)} |f(j, p_j + u_j) - f(j, p_j + v_j)|$$

$$\le \frac{1}{(m-1)!}\sum_{j=n}^{\infty} j^{(m-1)} g_j |u_j - v_j| \le \frac{1}{(m-1)!}\sum_{j=n}^{\infty} j^{m-1} g_j q_j \sup_{t\ge n_0} \frac{|u_t - v_t|}{q_t}$$

$$= ||u-v|| \frac{1}{(m-1)!}\sum_{j=n}^{\infty} j^{m-1} g_j q_j .$$

Hence by (8) and (9) we have

$$\sup_{n\ge n_0} \frac{|\hat{u}_n - \hat{v}_n|}{q_n} \le \frac{||u-v||}{(m-1)!} \cdot \sup_{n\ge n_3} \frac{1}{q_n}\sum_{j=n}^{\infty} j^{m-1} g_j q_j .$$

$$\le ||u-v|| \frac{c_1 + (\varepsilon/2M)}{(m-1)!} = \theta ||u-v||$$

where $\theta \in (0,1)$.

The above inequality yields

$$||\hat{u}-\hat{v}|| \le \theta ||u-v||,$$

where $\theta \in (0,1)$. Consequently, there is $w \in m_M(q)$ such that $w = Tw$. We find the equation for which w is a solution and next we find the solution x of (E) by means of w. Since the series $\hat{h}_n$ and $\hat{h}_{n+1}$ converge so do the series $\Delta \hat{h}_n = \hat{h}_{n+1} - \hat{h}_n$ and

$$\Delta \hat{h}_n = (-1)^{m-1} \sum_{j=n+1}^{\infty} \frac{(j+m-1-n-1)^{(m-1)}}{(m-1)!} f(j, p_j + h_j)$$

$$-(-1)^{m-1} \sum_{j=n}^{\infty} \frac{(j+m-1-n)^{(m-1)}}{(m-1)!} f(j, p_j + h_j)$$

$$= (-1)^{m-1}\left[\sum_{j=n}^{\infty} \frac{(j+m-1-n-1)^{(m-1)}}{(m-1)!} f(j, p_j + h_j) \right.$$

$$\left. - \sum_{j=n}^{\infty} \frac{(j+m-1-n)^{(m-1)}}{(m-1)!} f(j, p_j + h_j) \right]$$

$$= (-1)^{m} \sum_{j=n}^{\infty} \frac{(j+m-1-n)^{(m-1)} - (j+m-2-n)^{(m-1)}}{(m-1)!} f(j, p_j + h_j)$$

$$= (-1)^{m} \sum_{j=n}^{\infty} \frac{(j+m-2-n)^{(m-2)}}{(m-2)!} f(j, p_j + h_j).$$

Repeating the above reasoning we obtain convergence of $\Delta^r \hat{h}_n$ for $0 \le r \le m-1$ and

$$\Delta^r \hat{h}_n = (-1)^{m-1+r} \sum_{j=n}^{\infty} \frac{(j+m-1-n-r)^{(m-r-1)}}{(m-r-1)!} \times$$

$$f(j, p_j+h_j), \ 0 \leq r \leq m-1. \tag{24}$$

For $r = m-1$, (24) is

$$\Delta^{m-1} \hat{h}_n = \sum_{j=n}^{\infty} f(j, p_j+h_j)$$

We thus get

$$\Delta^m \hat{h}_n = \sum_{j=n+1}^{\infty} f(j, p_j+h_j) - \sum_{j=n}^{\infty} f(j, p_j+h_j)$$

$$= -f(n, p_n+h_n). \tag{25}$$

For $h = w$ we get from (25)

$$\Delta^m w_n = -f(n, p_n+w_n). \tag{26}$$

Setting $x_n = P_n + w_n$, recalling $\Delta^m p_n = 0$, (26) yields

$$\Delta^m x_n = -f(n, x_n)$$

i.e. x is the solution of (E) for $n \geq n_3$. To conclude the proof, we shall concentrate on relation (AR). For this we need any bound of $\Delta^r w_n$. From (24) and (18)

$$\Delta^r w_n = (-1)^{m+r} \sum_{j=n}^{\infty} \frac{(j+m-1-n-r)^{(m-r-1)}}{(m-r-1)!} \frac{\Delta I(j,w)}{j^{m-1}}$$

$$= \frac{(-1)^{m+r}}{(m-r-1)!} \sum_{j=n}^{\infty} \frac{(j+m-1-n-r)^{(m-r-1)}}{j^{m-r-1}} j^{-r} \Delta I(j,w)$$

$$= \frac{(-1)^{m+r}}{(m-r-1)!} \sum_{j=n}^{\infty} j^{-r} \left[\prod_{k=1}^{m-r-1} \left(1 + \frac{k-n}{j}\right) \right] \Delta I(j,w).$$

Since

$$\sum_{j=n}^{t} j^{-r} \left[\prod_{k=1}^{m-r-1} \left(1+\frac{k-n}{j}\right) \right] \Delta I(j,w)$$

$$= t^{-r} \left[\prod_{k=1}^{m-r-1} \left(1+\frac{k-n}{t}\right) \right] I(t+1,w)$$

$$- n^{-r} \prod_{k=1}^{m-r-1} \left(1+\frac{k-n}{n}\right) I(n,w)$$

$$-\sum_{j=n}^{t-1} I(j+1,w)\,\Delta\left[j^{-r}\prod_{k=1}^{m-r-1}\left(1+\frac{k-n}{j}\right)\right].$$

By using arguments similar to (17) we get

$$\Delta^r w_n = \frac{(-1)^{m+r}}{(m-r-1)!}\left\{-n^{-r}\frac{(m-r-1)!}{n^{m-r-1}}I(n,w)\right.$$

$$\left.-\sum_{j=n}^{\infty} I(j+1,w)\,\Delta\left[j^{-r}\prod_{k=1}^{m-r-1}\left(1+\frac{k-n}{j}\right)\right]\right\}.$$

$$=\frac{(-1)^{m+r+1}}{(m-r-1)!}\left\{\frac{(m-r-1)!}{n^{m-1}}I(n,w)\right.$$

$$+\sum_{j=n}^{\infty} I(j+1,w)\left[(j+1)^{-r}\Delta\prod_{k=1}^{m-r-1}\left(1+\frac{k-n}{j}\right)\right.$$

$$\left.\left.+\left(\prod_{k=1}^{m-r-1}\left(1+\frac{k-n}{j}\right)\right)\Delta j^{-r}\right]\right\}.$$

Hence, by virtue of (21)

$$|\Delta^r w_n| \le \frac{1}{(m-r-1)!}\left\{\frac{(m-r-1)!}{n^{m-1}}|I(n,w)|\right.$$

$$+\sum_{j=n}^{\infty}|I(j+1,w)|\left|\left[(j+1)^{-r}\Delta\prod_{k=1}^{m-r-1}\left(1+\frac{k-n}{j}\right)\right.\right.$$

$$\left.\left.\left.+\left(\prod_{k=1}^{m-r-1}\left(1+\frac{k-n}{j}\right)\right)\Delta j^{-r}\right]\right|\right\}$$

$$\le\frac{1}{(m-r-1)!}\left\{\frac{(m-r-1)!}{n^{m-1}}q_n s_n\right.$$

$$+q_n s_n\sum_{j=n}^{\infty}(j+1)^{-r}\Delta\prod_{k=1}^{m-r-1}\left(1+\frac{k-n}{j}\right)$$

$$\left.\left.+\left(\prod_{k=1}^{m-r-1}\left(1+\frac{k-n}{j}\right)\right)\Delta\left(-j^{-r}\right)\right]\right\}.$$

But

$$\lim_{t\to\infty}\left\{\sum_{j=n}^{t}(j+1)^{-r}\Delta\prod_{k=1}^{m-r-1}\left(1+\frac{k-n}{j}\right)\right.$$

$$+\sum_{j=n}^{t}\left[\prod_{k=1}^{m-r-1}\left(1+\frac{k-n}{j}\right)\right]\Delta\left(-j^{-r}\right)\Bigg\}$$

$$\leq \lim_{t\to\infty}\left\{\frac{1}{(n+1)^r}\left[\prod_{k=1}^{m-r-1}\left(1+\frac{k-n}{t+1}\right)-\prod_{k=1}^{m-r-1}\left(1+\frac{k-n}{n}\right)\right]\right.$$

$$\left.+\sum_{j=n}^{t}\Delta\left(-j^{-r}\right)\right\}=\frac{1}{(n+1)^r}\left[1-\frac{(m-r-1)!}{n^{m-r-1}}\right]+\frac{1}{(n+1)^r}$$

$$\leq\frac{1}{n^r}\left[1-\frac{(m-r-1)!}{n^{m-r-1}}\right]+\frac{1}{n^r}=\frac{2}{n^r}-\frac{(m-r-1)!}{n^{m-1}};$$

therefore

$$|\Delta^r w_n|\leq\frac{1}{(m-r-1)!}\left[\frac{(m-r-1)!}{n^{m-1}}q_n s_n+\frac{2}{n^r}q_n s_n\right.$$

$$\left.-\frac{(m-r-1)!}{n^{m-1}}q_n s_n\right]=\frac{2}{m-r-1!}\frac{q_n s_n}{n^r}\,. \quad (2$$

For $r=m-1$, $\Delta^{m-1}w_n$ is estimated as follows

$$|\Delta^{m-1}w_n|=\left|\sum_{j=n}^{\infty}\frac{\Delta I(j,w)}{j^{m-1}}\right|\leq\frac{2}{n^{m-1}}q_n s_n. \quad (2$$

Equation (27) which hold for $O<r<m-1$, (28) together with (23) gives

$$|\Delta^r w_n|\leq\frac{2(m-1)!M}{(m-r-1)!}q_n n^{-r},\ O<r\leq m-1. \quad ($$

Recalling $x_n=p_n+w_n$, by (23), and (29) we obtain

$$\Delta^r x_n=\Delta^r p_n+O(n^{-r}q_n),\ O\leq r\leq m-1.$$

Q.E

3. Remarks and examples

Remark 1. We have noticed that a solution x of (E), for which asymptotic rela (AR) holds, exists for n sufficiently large. For many applications we need solution for all $n\in N$. We rewrite equation (E) in the equivalent form

$$(-1)^m y_n+f(n,y_n)=-\sum_{i=1}^{m}\binom{m}{i}(-1)^{m-i}y_{n+i}.$$

From this it is evident that if the functions z_n defined by

$$z_n(x):=(-1)^m x+f(n,x),\ n_0\leq n<n_3$$

are surjections of R onto R, then there exist a solution x of (E) defined for all $n \in N$ which satisfies the asymptotic relation (AR). To this we put in the equality (30) $n = n_3 - 1, y_{n+i} = x_{n_3+i-1}$, $i = 1, \ldots, m$, whose existence is ensured by the theorem. The solution may not be uniquely denoted by x_{n_3-1}. Repeating this way we find in succession $x_{n_3-2}, \ldots, x_{n_0}$. This solution satisfies (AR).

Remark 2. If $c_1 = c_2 = 0$, then by (20) $\lim_{n\to\infty} s_n = 0$. Therefore from (23), (27), and (28) we have $\lim_{n\to\infty} \Delta^r w_n = 0$, $0 \le r \le m-1$. Hence the solution x of (E) satisfies the asymptotic relation

$$\Delta^r x_n = \Delta^r p_n + o(n^{-r} q_n). \tag{AR1}$$

Remark 3. If instead of (2) we have

$$\sum_{j=n_0}^{\infty} j^{m-1} g_j < \infty \tag{31}$$

then (3) holds for any nonincreasing q, with $c_1 = 0$. If (2) holds and $\lim_{n\to\infty} q_n \neq 0$ then (31) is fulfilled. Furthermore by (4) it follows from (5) that $c_2 = 0$. So in this case the solution x has the asymptotic behaviour (AR1).

Remark 4. In some distances we may take M as arbitrarily large. Then (6) is no restriction.

Example. Suppose

$$\sum_{j=n_0}^{\infty} j^{2m-2} a_j \text{ and } \sum_{j=n_0}^{\infty} j^{m-1} b_j \tag{32}$$

converge—perhaps conditionally—and

$$\sum_{j=n_0}^{\infty} j^{m-1} |a_j| \tag{33}$$

converges. Let p be any polynomial of degree $< m$. Then the theorem implies that the equation

$$\Delta^m y_n + a_n y_n = b_n \tag{E1}$$

has a solution x such that

$$\Delta^r x_n = \Delta^r p_n + o(n^{-r}), \; 0 \le r \le m-1. \tag{34}$$

To see this take $q \equiv 1$ and $f(n,u) = a_n u - b_n$. Then (1) holds with $g_n = |a_n|$. Hence by (32), (33) the assumptions (2) to (6) are satisfied with $c_1 = c_2 = 0$. Therefore, by Remark 2 we get the existence of a solution x of (E1) for which (34) holds.

As the second example we consider the equation

$$\Delta^m y_n + a_n (y_n)^s = b_n, \; s \geq 1. \tag{E2}$$

If

$$\sum_{j=n_0}^{\infty} j^{m-1} |a_j| \text{ and } \sum_{j=n_0}^{\infty} j^{m-1} b_j \tag{35}$$

then by the theorem there exists a solution x of (E2) which has an asymptotic property

$$x_n = c + o(1)$$

for arbitrary positive constant c. In this case we can take $q \equiv 1$, $g_n = s(M+c)^{s-1}|a_n|$ and apply Remark 2 with $r = 0$, $p = c$.

Remark 5. Based on the above two examples we compare the result of this note with the other contained in earlier papers. Equation (E2) cannot be studied by Theorem 3.1 [2] because of some kind of submultiplicity assumptions for the function f contained there and by theorems contained in [8] because as it was noticed in this paper the case $s > 1$ was ruled out. Asymptotic properties of the solutions of equations (E1) and (E2) can be studied by Theorem 3.6 of [1]; however this theorem gives conditions under which the solution x is asymptotic to the polynomial of degree exactly $m-1$. Instead of convergence of the second series (35) it is supposed that the series

$$\sum_{j=n_0}^{\infty} |b_j| \tag{36}$$

converges. Examining equation (E1) by Theorem 2 of [6] we state that if the series (36) and (33) converge every solution of (E1) satisfies asymptotic relation

$$\Delta^r x_n = O(n^{m-r-1}), \; 0 \leq r \leq m-1.$$

By Theorem 3 of [6], if both the series of (32) converge absolutely, then every solution of (E1) is of the form

$$x_n = p_n + o(1),$$

where the polynomial p_n has calculated coefficients. The theorem presented in [7] gives conditions when the solutions are of the form such that

$$\lim_{n \to \infty} \frac{\Delta^r x_n}{\Delta^r z_n} = c, \; 0 \leq r \leq m-1,$$

where z is such that

$$\lim_{n \to \infty} \Delta^{m-1} z_n = \infty$$

and so z_n cannot be polynomial of degree $< m$. Note that the equations considered in [1], [2], [6], [7] are of general form than (E).

References

[1] Agarwal R P, Properties of solutions of higher order nonlinear difference equations II, *Ann. Stii. Univ. Iasi.* **24** (1983) 85–96

[2] Agarwal R P, Properties of solutions of higher order nonlinear difference equations, *Ann. Stii. Univ. Iasi.* (in press)

[3] Drozdowicz A and Popenda J, Asymptotic behaviour of the solutions of the second order difference equation, *Proc. Am. Math. Soc.* (in press)

[4] Hooker J W and Patula W T, A second order nonlinear difference equation: oscillation and asymptotic behaviour, *J. Math. Anal. Appl.* **91** (1983) 9–29

[5] Patula W T, Growth, oscillation and comparison theorems for second order linear difference equations, *SIAM J. Math. Anal.* **10** (1979) 1272–1279

[6] Popenda J, On the asymptotic behaviour of the solutions of an *n*th order difference equation, *Ann. Polon. Math.* **44** (1984) 95–111

[7] Popenda J, On the asymptotic behaviour of solutions of difference equations, *Ann. Polon. Math.* (in press)

[8] Popenda J and Schmeidel E, On the asymptotic behaviour of non-oscillatory solutions of difference equations, *Fasc. Math.* **12** (1980) 43–53

[9] Trench W F, Asymptotic behaviour of solutions of an *n*th order differential equation, *Rocky Mountain J. Math.* **14** (1984) 441–450

SUBJECT INDEX

AUTHOR INDEX